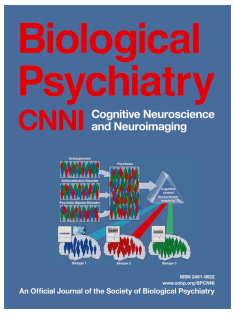
DOI:
10.1016/j.bpsc.2018.11.013

Document Version
Peer reviewed version

Accepted Manuscript



Dissecting the heterogeneous cortical anatomy of autism spectrum disorder using normative models.

Mariam Zabihi, Marianne Oldehinkel, Thomas Wolfers, Vincent Frouin, David Goyard, Eva Loth, Tony Charman, Julian Tillmann, Tobias Banaschewski, Guillaume Dumas, Rosemary Holt, Simon Baron-Cohen, Sarah Durston, Sven Bölte, Declan Murphy, Christine Ecker, Jan K. Buitelaar, Christian F. Beckmann, Andre F. Marquand

PII: S2451-9022(18)30329-X

DOI: https://doi.org/10.1016/j.bpsc.2018.11.013

Reference: BPSC 369

To appear in: *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*

Received Date: 18 October 2018

Revised Date: 30 November 2018

Accepted Date: 30 November 2018

Please cite this article as: Zabihi M., Oldehinkel M., Wolfers T., Frouin V., Goyard D., Loth E., Charman T., Tillmann J., Banaschewski T., Dumas G., Holt R., Baron-Cohen S., Durston S., Bölte S., Murphy D., Ecker C., Buitelaar J.K., Beckmann C.F. & Marquand A.F., Dissecting the heterogeneous cortical anatomy of autism spectrum disorder using normative models., *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* (2019), doi: https://doi.org/10.1016/j.bpsc.2018.11.013.

Dissecting the heterogeneous cortical anatomy of autism spectrum disorder using normative models.

Mariam Zabihi[1,2*], Marianne Oldehinkel[1,2], Thomas Wolfers[3,2], Vincent Frouin[4], David Goyard[4], Eva Loth[5], Tony Charman[6], Julian Tillmann[6], Tobias Banaschewski[7], Guillaume Dumas[8], Rosemary Holt[9], Simon Baron-Cohen[9], Sarah Durston[10], Sven Bölte[11,12], Declan Murphy[13,5], Christine Ecker[14,5], Jan K. Buitelaar[1,2,15], Christian F. Beckmann[1,2,16], Andre F. Marquand[1,2,17]

[1] Department of Cognitive Neuroscience, Radboud University Medical Center, Nijmegen, The Netherlands

[2] Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands

[3] Department of Human Genetics, Radboud University Medical Center, Nijmegen, The Netherlands

[4] Neurospin, Institut Joliot, CEA, Université Paris-Saclay, Gif-sur-Yvette, France

[5] Department of Forensic and Neurodevelopmental Sciences, Institute of Psychiatry, Psychology and Neuroscience King's College London, United Kingdom.

[6] Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience King's College London, United Kingdom.

[7] Department of Child and Adolescent Psychiatry and Psychotherapy, Central Institute of Mental Health Mannheim, Mannheim, Germany

[8] Human Genetics and Cognitive Functions Unit, Institut Pasteur, 25 Rue du Docteur Roux, Paris Cedex 15, France.

[9] Autism Research Centre, Department of Psychiatry, University of Cambridge, United Kingdom

[10] University Medical Centre, Utrecht, Utrecht, The Netherlands

[11] Center for Neurodevelopmental Disorders (KIND), Division of Neuropsychiatry, Department of Women's and Children's Health, Stockholm, Sweden

[12] Child and Adolescent Psychiatry, Centre of Psychiatry Research, Stockholm County Council, Sweden

[13] Sackler Institute for Translational Neurodevelopment, Institute of Psychiatry, Psychology and Neuroscience, King's College London, United Kingdom

[14] Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital Frankfurt am Main, Goethe University

[15] Karakter Child and Adolescent Psychiatry University Centre, Nijmegen, The Netherlands

[16] Centre for Functional MRI of the Brain (FMRIB), University of Oxford, Oxford, United Kingdom

[17] Department of Neuroimaging, Institute of Psychiatry, Psychology and Neuroscience King's College London, United Kingdom.

corresponding author: Mariam Zabihi,
Address: Kapittelweg 29, 6525 EN Nijmegen, Gelderland, Netherlands, Phone: +31- 243668494,
Email: m.zabihi@donders.ru.nl

Short title: Dissecting heterogeneity of ASD with normative modeling

Keywords: Normative modeling; Gaussian process; Autism; cortical thickness; Outlier detection; heterogeneity

Word count of the abstract: 249
Word count of the main text: 4000
Number of tables: 1
Number of figures: 8
Number of supplementary material: 1

Abstract

Background

The neuroanatomical basis of autism spectrum disorder (ASD) has remained elusive, mostly due to high biological and clinical heterogeneity among diagnosed individuals. Despite considerable effort towards understanding ASD using neuroimaging biomarkers, heterogeneity remains a barrier, partly because studies mostly employ case-control approaches, which assume that the clinical group is homogeneous.

Methods

Here, we used an innovative normative modelling approach to parse biological heterogeneity in ASD. We aimed to dissect the neuroanatomy of ASD by mapping the deviations from a typical pattern of neuroanatomical development at the level of the individual and to show the necessity to look beyond the case-control paradigm to understand the neurobiology of ASD. We first estimated a vertex-wise normative model of cortical thickness development using Gaussian process regression, then mapped the deviation of each participant from the typical pattern. For this we employed a heterogeneous cross-sectional sample of 206 typically developing (TD) individuals (127 male), and 321 individuals (232 male) with ASD (aged 6-31).

Results

We found few case-control differences but the ASD cohort showed highly individualized patterns of deviations in cortical thickness that were widespread across the brain. These deviations correlated with severity of repetitive behaviors and social communicative symptoms, although only repetitive behaviors survived corrections for multiple testing.

Conclusions

Our results: (i) reinforce the notion that individuals with ASD show distinct, highly individualized trajectories of brain development and (ii) show that by focusing on common effects (i.e. the 'average ASD participant'), the case-control approach disguises considerable inter-individual variation crucial for precision medicine.

Introduction

Autism spectrum disorder (ASD) is a lifelong neurodevelopmental disorder diagnosed exclusively on the basis of symptomatology, period of onset, and impairment (i.e., impairments in social-communication and interaction, alongside repetitive stereotyped behavior and sensory anomalies)(1). Autism is well recognized as being highly heterogeneous on multiple levels: for example, in terms of its clinical presentation and underlying neurobiology. Indeed, more than 100 genes(2) and many aspects of brain structure have been associated with ASD at the group level(3). Autism is also grounded in the process of brain maturation and it is believed that alterations are evident throughout brain development(4, 5). In particular, differences in cortical thickness (CT) have been reported across different studies and ages(6), which – together with differences in surface area (SA)(6–10) – underpin regional differences in brain volume in ASD(10–13). However, the precise etiology of the disorder in terms of brain development and underlying mechanisms remain elusive.

The heterogeneity of ASD is a fundamental barrier to understanding the neurobiology of ASD and the development of interventions(14). Regional group-level differences have been reported across several neuroanatomical measures, including cortical thickness (8, 10, 15–22). However these findings show generally poor replication across studies(3, 7, 19, 23, 24) and small effect sizes(8, 19). Heterogeneity is also evident in studies that have used classifiers to discriminate ASD participants from controls, which mostly show relatively low accuracy for predicting diagnosis, especially in large samples(19, 25, 26). An important reason for this is that most studies to date have employed a traditional case-control approach, which is based on the assumption that the clinical and control groups are homogeneous entities(7, 27). Thus, the case-control approach provides information about alterations at the group level or, in other words, in the 'average ASD participant'. However, different participants may have different symptom profiles and different etiological pathways and resulting neurobiological changes may converge on the same symptoms. Therefore, to understand the neurobiology of ASD, it is important to understand the range of associated neurobiological variation, which may subsequently inform intervention at the level of the individual in the spirit of 'precision

medicine'(28). A common approach to study the biological heterogeneity underlying ASD is to find subtypes using clustering algorithms, mostly on the basis of symptoms or behavioral characteristics(29–34). This approach has been somewhat successful and is appropriate if the clinical cohort can be cleanly partitioned into a relatively small number of homogeneous subgroups on the basis of the chosen measures. However, it does not tackle heterogeneity within subgroups and it may be the case that no clearly defined subgroups exist in the data. Moreover, subgroups derived from behavior or symptoms require extensive validation on external measures and still may not fully reflect the underlying biology(35, 36).

Here, we apply a complementary normative modelling approach(36, 37) to understand the biological heterogeneity of ASD. This shifts the focus away from group-level comparisons – which can detect consistent differences across groups of individuals (e.g. diagnoses or putative subtypes) – towards characterizing the degree of alteration in each individual, with reference to the typically developing brain. This allows us to detect and map neuroanatomical alterations at the level of the individual and has recently shown promise in understanding the biological variation of psychotic disorders(37) . Normative modelling is analogous to the use of growth charts in pediatric medicine, which allow the development (e.g. in terms of height or weight) of each individual child to be measured against expected centiles of variation in the population. To achieve this, we first estimated a statistical model characterizing typical cortical development that accurately quantifies the variation within the population and across brain development. We then placed each individual ASD participant in relation to the typical distribution in order to identify alterations in individual cases with respect to the typical pattern of brain maturation. Our main goals were to: (ii) to map the neuroanatomical features by which each individual ASD participant differs from the expected typically developing pattern, across both different developmental stages and levels of functioning and thereby (ii) demonstrate the value of normative modelling techniques for understanding the biological heterogeneity of ASD. For this, we employed data from a large international study(38) with harmonized data acquisition procedures and a design that naturally groups of subjects according to

different developmental stages. Whilst normative modelling is suitable for many different aspects of brain structure or function, here, we focused on cortical thickness which is a sensitive and reliable measure of cortical morphology in ASD(6, 8, 39), although we also investigated surface area. Ultimately, we hope this approach will yield a set of individualized neurobiological 'fingerprints' facilitating a route towards precision medicine approaches in ASD(28).

Materials and Methods

Participants

Full details on study design and clinical characteristics have been described previously(38). Briefly, we included all participants from the Longitudinal European Autism Project (LEAP)(40) cohort with a structural MRI scan surviving quality control and the necessary clinical and demographic data. We included 206 typically developing (TD) individuals aged 7 to31 years (127 male; Table 1; Table S1, Figure S1) and 321 individuals aged 6 to31 years (232 male) with ASD. There were no significant differences between TD and ASD cohorts in age but the IQ of ASD participants was lower than TD participants. Under the study design, each cohort was split into four subgroups according to age and level of intellectual ability (Table 1): (i) Adults with ASD without intellectual disability (ID) and TD controls aged 18 to 30 years, IQ ≥ 70; (ii) Adolescents with ASD without ID and TD aged 12 to 17 years; (iii) Children with ASD without ID or TD aged 6 to 11 years; and (iv) Adolescents and adults with ASD and ID (i.e. full-scale IQ between 50 to 70(1)) aged 12 to 30 years. Note that only TD participants were included in the estimation of the normative model.

[Table 1]

TD participants were recruited via advertisement. Individuals with an existing ASD and/or mild ID diagnosis (according to DSM5/ICD10 criteria) were recruited from existing databases and clinic contacts across one of seven study sites: the Institute of Psychiatry, Psychology and Neuroscience, King's College London, UK, Autism Research Centre at the University of Cambridge, UK, Radboud University Nijmegen Medical Centre, University Medical Centre Utrecht, the Netherlands, Central

Institute of Mental Health, Mannheim, Germany, and the University Campus Bio-Medico, Rome, Italy. The combined information of Autism Diagnostic Interview-Revised(41) (ADI-R) and Autism Diagnostic Observation Schedule Second Edition(42) (ADOS-2) were used to measure symptom severity(33). However, individuals with a clinical ASD diagnosis who did not reach conventional cut-offs on these instruments were not excluded. The ADI-R is a parent reported measure of lifetime or past developmental window symptom severity whereas the ADOS-2 is an expert rating of current symptoms. A standard set of exclusion criteria were applied and are provided in the supplementary material. All subjects were scanned with a T1-weighted imaging protocol and Freesurfer (version 5.3) was used to estimate measures of regional CT and SA. See supplementary methods for details.

Constructing a normative model of cortical thickness

An overview of the normative modelling approach is shown in Figure **1** and has been described previously(36). Briefly, Gaussian process regression (GPR)(43) was used to estimate separate normative models of CT and SA at each vertex on the cortical surface (see supplementary methods for details). This normative model can be used to predict both the expected CT and the associated predictive uncertainty for each individual participant. The contours of predictive uncertainty can then be used to model centiles of variation within the cohort. This allows us to place each individual participant within the normative distribution thereby quantifying the vertex-wise deviation of CT from the healthy range across the brain.

*[*Figure 1*]*

To achieve this, we generated a developmental model of typical brain development by training a GPR model on the TD cohort (N=206) using age and gender as covariates (i.e. independent variables) to predict cortical thickness (i.e. dependent variable). In pediatric medicine, growth charts are normally estimated on the basis of a large population cohort (i.e. potentially including patients with various disorders based on the population prevalence). In our sample, the prevalence of ASD is much higher

than in the population, so for simplicity and to avoid the normative model being enriched for ASD, we estimated the normative model on the basis of the TD participants only. Moreover, while the number of data we employ here is relatively small in comparison with population based studies, our Bayesian statistical model provides a principled method to handle uncertainty and therefore automatically makes inferences more conservative as the number of data points decreases, although more data would allow more precise estimates. To assess generalization, we used 10-fold cross-validation before retraining the model using the whole dataset to make predictions on the ASD participants following standard practice in machine learning (supplementary methods). Importantly, all parameters were estimated using the training data using empirical Bayesian estimation(36) and the use of cross-validation ensures unbiased estimates for the TD cohort as well as for the ASD cohort. Therefore, deviations can be compared with one another.

Estimating regional deviations for each subject

To estimate a pattern of regional deviations from typical cortical thickness for each participant, we derived a normative probability map (NPM) that quantifies the deviation from the normative model for cortical thickness at each vertex. This was done by using the normative model to predict vertex-wise estimates of cortical thickness for each individual participant, then estimating a subject-specific Z-score(36) (supplementary methods). This provides a statistical estimate of how much each individual differs from the healthy pattern at each vertex. We thresholded the NPMs, correcting for multiple comparisons by controlling false discovery rate (FDR) at $p < 0.05$ within each participant, as in(36).

To measure the spatial overlap of the individualized deviations across the cohort, we calculated an overlap map by counting the significant (FDR corrected) vertices derived from the Z-score maps across all subject-level NPMs. The resulting summary maps indicate the spread of vertex-wise deviations across the brain, separately for positive and negative deviations. This allowed us to

identify a set of brain regions where participants had increased (positive deviation) or decreased (negative deviation) cortical thickness relative to the reference cohort.

To provide a simple comparison for these subject-level deviations, we also estimated a standard vertex-wise general linear model to establish significant differences between groups including age as a covariate. We also investigated models including quadratic and cubic age terms (corrected using FDR at $p < 0.05$) and separate models for males and females.

Constructing an individual-level atypicality score

A key benefit of normative modelling is a probabilistic interpretation of the deviations across all subjects. The NPMs therefore provide a multivariate measure of deviation from the normative range across all brain regions. This captures spatially distributed differences from the TD pattern. To better understand most important focal differences for each subject we estimated a summary score for each participant capturing the individual's largest deviation from the typical pattern (which is potentially the most clinically relevant). This can be modelled using extreme value statistics (44) and is based on the notion that the expected maximum of any random variable converges to an extreme value distribution (EVD). Therefore, we estimated a maximum deviation for each subject by taking a trimmed mean of 1% of the top absolute deviations for each subject across all vertices and fit an EVD to these deviations.

Mapping behavioral associations

Last, to assess the clinical relevance of these deviations, we computed Spearman correlation coefficients between global and regional extreme deviation from the normative model and ADOS-2/ADI-R symptom severity scores ($p < 0.05$, FDR). The global measure (described above) provides an overall summary of the deviation for each individual whilst the regional assessment helps to determine the functional correspondence of the deviations across individuals on a region-by-region

basis. The regional extreme deviation was computed as the trimmed mean of the 1% of top absolute deviations for each region after parcellating the cortex using the Desikan-Killiany atlas(45).

Checking for potential confounds

To investigate whether potential confounds could have influenced our findings, we estimated a separate normative model additionally including dummy regressors for IQ, site and Freesurfer Euler number (46). We also performed post-hoc tests between the deviations from the normative model and potential confounding variables (IQ, comorbid symptoms, and surrogate measures of image quality) See Tables S2 and S3.

Results

A normative model quantifying the decline of cortical thickness with age

Figure 2 shows the developmental normative model of CT derived from the TD male cohort, thresholded to show vertices where the correlation between true and predicted labels was higher than predicted by chance ($p < 0.05$, FDR corrected. See Figure S3 for females). The unthresholded map showing the correlation between true and predicted CT values is shown in Supplementary Figure S2 along with the root mean squared error of the normative model across different vertices. In most regions, CT decreases consistently and approximately linearly with age. However, in some regions, CT followed a non-linear (i.e. an inverted U-shaped) trajectory with an early rise followed by a decline, e.g. in the inferior temporal and posterior frontal regions. This corresponds well with the known developmental trajectory of CT(47–51). The normative model for SA showed a similar, relatively global pattern of decline as for CT (not shown).

*[*Figure 2*]*

Widespread deviations from normative pattern of cortical thickness among the ASD cohort

Figure 3 shows the classical mass-univariate group difference (i.e. case control) map between ASD and TD cohorts. This shows few significant differences between groups; only two small regions of increased CT in superior frontal and parietal cortices survived FDR correction. There were also few significant differences when additionally including quadratic and cubic age terms and no differences in the age x diagnosis interaction. The separate models for males and females also did not show any significant differences after FDR correction.

[Figure 3]

Figure 4 and Figure 5 show a summary of the NPMs for the ASD and TD cohort. Specifically, these figures show the number of participants in each group that deviate negatively (Figure 4) or positively (Figure 5) from the normative model at each vertex after intra-individual FDR correction. Importantly, and contrast to the GLM, these deviations need not overlap between subjects. As expected, the TD cohort shows few significant deviations, indicating that the normative model provides a good fit for this cohort. Crucially, this fit was achieved under cross-validation and is therefore unbiased. Therefore, under the 'null' hypothesis that ASD participants follow a similar trajectory of brain development to TD participants there is no prior reason to expect the fit will be better in TD than in ASD participants. In contrast, the total number of deviating vertices was noticeably higher in the ASD cohort and was widespread across the brain, suggesting that there are widespread and individualized deviations from the normative model in certain subsets of participants. When considering each age group separately, negative deviations were most prominent in children, whereas positive deviations were most prominent in adolescents and adults. The results were very similar for the models including IQ, scanning site and Euler number as covariates (Figures S4 and S5) and a similar pattern of results was observed for SA, albeit with slight differences with respect to the pattern of deviations across brain regions (Figure S6 and S7).

[Figure 4]

[Figure 5]

ASD participants deviate more than TD from the normative pattern of development

Figure 6 shows the distribution of the most extreme deviations from the normative model across the brain. This shows that the maximum deviation across the brain is higher in the ASD cohort relative to the TD cohort and the distribution of the ASD cohort is shifted towards the right implying relatively more subjects with extreme deviations. Saliently, the top fifteen deviating individuals belong to the ASD cohort, which is extremely unlikely to occur by chance ($p < 0.0005$, binomial test). The NPMs of these participants (Figure S8) have highly individualized patterns of deviation not only with respect to brain regions but also in sign, with some participants having positive deviations (i.e., greater CT) or negative deviations (reduced CT). These participants did not show a consistent pattern with respect to their symptom scores (Table S5), which underscores the degree of clinical and neurobiological heterogeneity within the ASD cohort. However, with regard to their demographic profile, subjects with predominantly positive deviations were adolescents or adults, whilst most subjects with negative deviations were children.

*[*Figure 6*]*

Association with symptoms

Global deviations from the normative model were negatively associated with ADOS repetitive behaviours (ρ=-0.21 $p < 0.05$) and regional deviations were associated with symptoms in several brain regions (Figure 7 and Figure 8). Associations were found with symptom severity in the repetitive domain of the ADOS-2 or ADI-R in prefrontal regions in females. In males, a similar pattern was seen, but did not survive multiple comparison correction except the superiorfrontal region in ADI-R. Social interaction and communication scores also had nominally significant associations in females but these did not survive correction.

*[*Figure 7*]*

*[*Figure 8*]*

Discussion

In this study, we aimed to dissect the heterogeneous neurobiology of ASD by mapping the deviation of each individual participant from a normative model of cortical thickness development. In a large heterogeneous cohort spanning a wide range of the ASD phenotype, we showed few significant group-level differences between ASD and TD cohorts in cortical thickness using a classical case-control analysis. In contrast, our normative modelling approach showed striking, widespread patterns of cortical atypicality at the level of individual ASD participant. These patterns were highly individualized across participants, distinct across different developmental stages and associated with symptoms, especially repetitive behaviors. This supports the notion that a subset of ASD participants follow a different developmental trajectory to TD subjects, and that the trajectory each ASD participant follows is highly individualized. From a methodological standpoint, our study shows that: (i) that it is necessary to look beyond the case control paradigm to understand the heterogeneous neuroanatomy of ASD; (ii) that normative modelling provides as an alternative conceptual framework for understanding the heterogeneous neurobiology of ASD in terms of deviations from a typical pattern and (iii) that focusing on an 'average autistic individual' provides only a partial reflection of the nature of the condition. In other words, the case control approach focuses on common effects rather than inter-individual variation. Capturing and capitalizing on such variation at the individual level is at the heart of precision medicine.

The normative model describes the variation in typical brain development showed a largely monotonic –and in some areas non-linear– decrease of CT throughout development, consistent with previous neuroimaging studies(47–55). The fact that we observed widespread inter-individual differences between ASD participants in terms of their deviations from the normative model explains why our classical case-control analysis revealed few significant differences and why several large previous neuroimaging studies have also only detected relatively modest group level effects(8, 19). The heterogeneity underlying ASD is widely recognized(2, 56–62); some studies have reported reductions in CT in ASD(15) whereas some studies have reported increases(16, 63). Saliently, these inconsistencies remain evident even in large studies; for example, a large study derived from the

ENIGMA consortium demonstrated both regional increases and decreases in ASD at the group level that were consistent across development(8). Other studies – many derived from the ABIDE dataset(64) – have shown widespread increases in CT early in development that are attenuated later in development(19, 20, 48). Our results complement these studies because of our focus on studying individual variation within the ASD cohort. We show that: (i) a subset of participants show decreased CT and SA in childhood whilst (ii) other patients show regional increases in childhood in different areas (e.g. peri-calcarine cortex); (iii) some participants show increased CT and SA in adolescence or adulthood. Crucially however, these effects show minimal overlap across brain regions in different individuals. This is in line with another recent study applying normative modelling to ASD, which found effects in a subset of participants that were different from the main group effects(65). Thus, we consider that group-level effects can be understood as the background upon which individual variation is superimposed. The individualized deviations we report were mostly located in areas previously associated with ASD, such as the medial cortex including cingulate and dorsomedial prefrontal regions, lateral prefrontal and parietal cortices, temporal cortices and the hippocampal formation(6, 7, 63, 66, 67). Whilst some of these regions have been associated with social processing, the individual deviations in these regions were not associated with social interaction or communication symptoms at the group level. This could be for several reasons, for example the anatomical patterns associated with these symptoms may be expressed in other measures of cortical anatomy (e.g.(68, 69)) or in subcortical regions. Adults and adolescents had relatively fewer deviations, but these were positive (relatively increased CT and SA) and widespread across prefrontal and temporal cortices. Notably, we detected relatively few deviations in ASD with ID, which is important to exclude the possibility that these subjects were driving the effects described above. However, the ASD with ID group was relatively small (N = 20) so we do not draw strong conclusions about potential differences between ASD with- and without ID.

The 15 subjects with the most atypical anatomy all had ASD, which is extremely unlikely to occur by chance. Moreover, these participants had individualized brain alterations and clinical

characteristics. At the group level, the regional deviations we detected from the normative model were associated with the severity of lifetime and current autistic symptoms (ADI-R and ADOS-2 respectively), demonstrating that our model predictions may be clinically relevant. The deviation from the normative range was most informative about repetitive behavior symptom severity in that the strongest correlations were between CT in prefrontal regions with restricted repetitive behaviors, especially in females and across both parental report via ADI-R and observer ratings of current symptoms via ADOS-2. These results broadly correspond with previous reports(6, 70, 71) and suggest that ASD may be more heterogeneous in males, but we are cautious about this interpretation because we did not test it directly. Taken together, our results add weight to the importance of considering ASD in the context of a model of typical brain development and at the individual level(39, 63, 67)

Our findings should be considered in the light of several limitations. First, the trajectories of brain development were based on cross-sectional data and should be validated in a longitudinal cohort. Longitudinal follow-up data are currently being acquired and will be the subject of a future report. Moreover, while our sample size is similar to other neuroimaging studies of brain development (e.g. (72)), the model would yield more precise estimates with more data. Second, we registered all subjects to a standard adult template brain, as is standard in the field(10, 63, 67, 73–75), which could cause bias. However, there were few deviations in the TD cohort which makes this possibility unlikely. Third, our data does not permit strong inferences about the degree to which confounding variables may have influenced our findings. We found moderate associations between deviations from the normative models and a surrogate metric of image quality, but these were also associated with childhood ASD symptoms, comorbid ADHD symptoms and IQ. Moreover, our study design does not permit inferences about the direction of causality. For example, subjects with the most abnormal anatomy may also have the most impairment. Finally, we did not perform manual edits on the cortical surface reconstructions. Whilst this eliminates one potential source of bias, the

results need to be interpreted in the light of this and it is possible that performing manual edits may improve the quality of the surface reconstructions in some cases.

In conclusion, we estimated a normative model of cortical development based on a large typically developing cohort and applied this model to a heterogeneous ASD cohort. Our results show that it is necessary to look beyond the case-control paradigm –which is limited to detecting group-level effects describing the 'average ASD participant'– to understand the heterogeneous neurobiology of ASD. Normative modelling is well suited for this purpose as it can chart the individualized deviation of each individual subject relative to the normative range, and hence provides an excellent tool for understanding the heterogeneity of psychiatric disorders.

Acknowledgments

We gratefully acknowledge the support of the EU-AIMS LEAP study team for data acquisition, quality control and preprocessed and also support from the Netherlands Organization for Scientific Research (NWO) through VIDI grants to AFM (Grant No. 016.156.415) and CFB (864.12.003). JKB received funding from the FP7 under Grant Nos. 602805 (AGGRESSOTYPE), 603016 (MATRICS), and 278948 (TACTICS) and from the European Community's Horizon 2020 Programme (H2020/2014-2020) under Grant Nos. 643051 (MiND) and 642996 (BRAINVIEW). We also gratefully acknowledge funding from the Wellcome Trust UK Strategic Award (098369/Z/12/Z). This work was supported by EU-AIMS (European Autism Interventions), which receives support from the Innovative Medicines Initiative Joint Undertaking under grant agreement no. 115300, the resources of which are composed of financial contributions from the European Union's Seventh Framework Programme (grant FP7/2007-2013), from the European Federation of Pharmaceutical Industries and Associations companies' in-kind contributions.

The preprint of the paper is available via bioRxiv. **doi:** https://doi.org/10.1101/477596

Conflict of Interest

JKB has been a consultant to, advisory board member of, and a speaker for Janssen Cilag BV, Eli Lilly, Shire, Lundbeck, Roche, and Servier. He is not an employee of any of these companies, and not a stock shareholder of any of these companies. He has no other financial or material support, including expert testimony, patents or royalties. CFB is director and shareholder in SBGNeuro Ltd. Sven Bölte discloses that he has in the last 5 years acted as an author, consultant or lecturer for Shire, Medice, Roche, Eli Lilly, Prima Psychiatry, GLGroup, System Analytic, Ability Partner, Kompetento, Expo Medica, and Prophase. He receives royalties for text books and diagnostic tools from Huber/Hogrefe, Kohlhammer and UTB. Tobias Banaschewski served in an advisory or consultancy role for Actelion, Hexal Pharma, Lilly, Lundbeck, Medice, Novartis, Shire. He received conference support or speaker's fee by Lilly, Medice, Novartis and Shire. He has been involved in clinical trials conducted by Shire & Viforpharma. He received royalities from Hogrefe, Kohlhammer, CIP Medien, Oxford University Press.The present work is unrelated to the above grants and relationships The other authors report no biomedical financial interests or potential conflicts of interest.

References

1. American Psychiatric Association, American Psychiatric Association. DSM-5 Task Force (2013): *Diagnostic and statistical manual of mental disorders : DSM-5. DSM-5*. .

2. Betancur C (2011): Etiological heterogeneity in autism spectrum disorders: More than 100 genetic and genomic disorders and still counting. *Brain Res*. 1380: 42–77.

3. Ecker C (2016): The neuroanatomy of autism spectrum disorder: An overview of structural neuroimaging findings and their translatability to the clinical setting. *Autism*. 1362361315627136-.

4. Walsh C a, Morrow EM, Rubenstein JLR (2008): Essay Autism and Brain Development. *October*. 135: 396–400.

5. Schumann CM, Bloss CS, Barnes CC, Wideman GM, Carper RA, Akshoomoff N, *et al.* (2010): Longitudinal Magnetic Resonance Imaging Study of Cortical Development through Early Childhood in Autism. *J Neurosci*. 30: 4419–4427.

6. Ecker C, Ginestet C, Feng Y, Johnston P, Lombardo M V, Lai M-C, *et al.* (2013): Brain Surface Anatomy in Adults With AutismThe Relationship Between Surface Area, Cortical Thickness, and Autistic SymptomsBrain Surface Anatomy in Adults With Autism. *JAMA Psychiatry*. 70: 59–70.

7. Hyde KL, Samson F, Evans AC, Mottron L (2010): Neuroanatomical differences in brain areas implicated in perceptual and other core features of autism revealed by cortical thickness analysis and voxel-based morphometry. *Hum Brain Mapp*. 31: 556–566.

8. van Rooij D, Anagnostou E, Arango C, Auzias G, Behrmann M, Busatto GF, *et al.* (2017): Cortical and Subcortical Brain Morphometry Differences Between Patients With Autism Spectrum Disorder and Healthy Individuals Across the Lifespan: Results From the ENIGMA ASD Working Group. *Am J Psychiatry*. appi.ajp.2017.1.

9. Ecker C, Bookheimer SY, Murphy DGM (2015): Neuroimaging in autism spectrum disorder: Brain structure and function across the lifespan. *Lancet Neurol*. 14: 1121–1134.

10. Hazlett HC, Poe M, Gerig G, Styner M, Chappell C, Smith RG, *et al.* (2011): Early Brain Overgrowth in Autism Associated with an Increase in Cortical Surface Area Before Age 2 years. *Arch Gen Psychiatry*. 68: 467–476.

11. Piven J, Arndt S, Bailey J, Havercamp S, Andreasen NC, Palmer P (1995): An MRI study of brain size in autism. *Am J Psychiatry*. .

12. Piven J, Arndt S, Bailey J, Andreasen N (1996): Regional brain enlargement in autism: a magnetic resonance imaging study. *J Am Acad Child Adolesc Psychiatry*. . doi: 10.1097/00004583-199604000-00020.

13. Hardan AY, Minshew NJ, Mallikarjuhn M, Keshavan MS (2001): Brain volume in autism. *J Child Neurol*. . doi: 10.1177/088307380101600607.

14. Lai MC, Lombardo M V., Chakrabarti B, Baron-Cohen S (2013): Subgrouping the Autism "Spectrum": Reflections on DSM-5. *PLoS Biol.* 11. doi: 10.1371/journal.pbio.1001544.

15. Hadjikhani N, Joseph RM, Snyder J, Tager-Flusberg H (2006): Anatomical differences in the mirror neuron system and social cognition network in autism. *Cereb Cortex*. 16: 1276–1282.

16. Mak-Fan KM, Taylor MJ, Roberts W, Lerch JP (2012): Measures of cortical grey matter structure and development in children with autism spectrum disorder. *J Autism Dev Disord*. 42: 419–427.

17. Wallace GL, Dankner N, Kenworthy L, Giedd JN, Martin A (2010): Age-related temporal and parietal cortical thinning in autism spectrum disorders. *Brain*. 133: 3745–3754.

18. Scheel C, Rotarska-Jagiela A, Schilbach L, Lehnhardt FG, Krug B, Vogeley K, Tepest R (2011): Imaging derived cortical thickness reduction in high-functioning autism: Key regions and temporal slope. *Neuroimage*. 58: 391–400.

19. Haar S, Berman S, Behrmann M, Dinstein I (2016): Anatomical Abnormalities in Autism? *Cereb Cortex*. 26: 1440–1452.

20. Khundrakpam BS, Lewis JD, Kostopoulos P, Carbonell F, Evans AC (2017): Cortical Thickness Abnormalities in Autism Spectrum Disorders Through Late Childhood, Adolescence, and Adulthood: A Large-Scale MRI Study. *Cereb Cortex*. 27: 1721–1731.

21. Courchesne E, Karns CM, Davis HR, Ziccardi R, Carper R a, Tigue ZD, *et al.* (2011): Unusual brain growth patterns in early life in patients with autistic disorder: an MRI study. *Neurology*. 57: 2111–2111.

22. Ecker C, Shahidiani A, Feng Y, Daly E, Murphy C, D'Almeida V, *et al.* (2014): The effect of age, diagnosis, and their interaction on vertex-based measures of cortical thickness and surface area in autism spectrum disorder. *J Neural Transm*. 121: 1157–1170.

23. Ramaswami G, Geschwind DH (2018): Genetics of autism spectrum disorder. *HandbClinNeurol*. .

24. Zhang W, Groen W, Mennes M, Greven C, Buitelaar J, Rommelse N (2017): Revisiting subcortical brain volume correlates of autism in the ABIDE dataset: effects of age and sex. *Psychol Med*. c: 1–15.

25. Wolfers T, Buitelaar JK, Beckmann CF, Franke B, Marquand AF (2015): From estimating activation locality to predicting disorder: A review of pattern recognition for neuroimaging-based psychiatric diagnostics. *Neurosci Biobehav Rev*. 57: 328–349.

26. Sabuncu MR, Konukoglu E (2014): Clinical Prediction from Structural Brain MRI Scans: A Large-Scale Empirical Study. *Neuroinformatics*. 13: 31–46.

27. Damiano CR, Mazefsky CA, White SW, Dichter GS (2014): Future Directions for Research in Autism Spectrum Disorders. *J Clin Child Adolesc Psychol*. 43: 828–843.

28. Insel TR, Cuthbert BN (2015): Brain disorders? Precisely. *Science (80-)*. 348: 499–500.

29. Lombardo M V., Pierce K, Eyler LT, CarterBarnes C, Ahrens-Barbeau C, Solso S, *et al.* (2015): Different functional neural substrates for good and poor language outcome in autism. *Neuron*. 86: 567–577.

30. Fountain C, Winter AS, Bearman PS (2012): Six Developmental Trajectories Characterize Children With Autism. *Pediatrics*. 129: e1112–e1120.

31. Fair DA, Bathula D, Nikolas MA, Nigg JT (2012): Distinct neuropsychological subgroups in typically developing youth inform heterogeneity in children with ADHD. *Proc Natl Acad Sci*. . doi: 10.1073/pnas.1115365109.

32. Costa Dias TG, Iyer SP, Carpenter SD, Cary RP, Wilson VB, Mitchel SH, *et al.* (2015): Characterizing heterogeneity in children with and without ADHD based on reward system connectivity. *Dev Cogn Neurosci*. . doi: 10.1016/j.dcn.2014.12.005.

33. van Loo HM, de Jonge P, Romeijn J-W, Kessler RC, Schoevers RA (2012): Data-driven subtypes of major depressive disorder: a systematic review. *BMC Med*. . doi: 10.1186/1741-7015-10-156.

34. Bell MD, Corbera S, Johannesen JK, Fiszdon JM, Wexler BE (2013): Social cognitive impairments and negative symptoms in schizophrenia: Are there subtypes with distinct functional correlates? *Schizophr Bull*. . doi: 10.1093/schbul/sbr125.

35. Marquand AF, Wolfers T, Mennes M, Buitelaar J, Beckmann CF (2016): Beyond Lumping and Splitting: A Review of Computational Approaches for Stratifying Psychiatric Disorders. *Biol Psychiatry Cogn Neurosci Neuroimaging*. 1: 433–447.

36. Marquand AF, Rezek I, Buitelaar J, Beckmann CF (2016): Understanding Heterogeneity in Clinical Cohorts Using Normative Models: Beyond Case-Control Studies. *Biol Psychiatry*. 80: 552–561.

37. Wolfers T, Doan N, Kaufmann T, al et (2018): Mapping the heterogeneous phenotype of schizophrenia and bipolar disorder using normative models. *JAMA Psychiatry*. . Retrieved from http://dx.doi.org/10.1001/jamapsychiatry.2018.2467.

38. Loth E, Charman T, Mason L, Tillmann J, Jones EJH, Wooldridge C, *et al.* (2017): The EU-AIMS Longitudinal European Autism Project (LEAP): design and methodologies to identify and validate stratification biomarkers for autism spectrum disorders. *Mol Autism*. 8: 24.

39. Anagnostou E, Taylor MJ (2011): Review of neuroimaging in autism spectrum disorders: What have we learned and where we go from here. *Mol Autism*. 2: 1–9.

40. Charman T, Loth E, Tillmann J, Crawley D, Wooldridge C, Goyard D, *et al.* (2017): The EU-AIMS Longitudinal European Autism Project (LEAP): clinical characterisation. *Mol Autism*. 8: 27.

41. Rutter M (2003): Autism diagnostic interview-revised. *West Psychol Serv*. .

42. Lord C, Risi S, Lambrecht L, Cook EH, Leventhal BL, Dilavore PC, *et al.* (2000): The Autism Diagnostic Observation Schedule-Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *J Autism Dev Disord*. . doi: 10.1023/A:1005592401947.

43. Rasmussen CE, Williams CKI (2006): Model Selection and Adaptation of Hyperparameters. *Gaussian Process Mach Learn (Adaptive Comput Mach Learn Ser*. 105–128.

44. Fisher RA, Tippett LHC (1928): Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Math Proc Cambridge Philos Soc*. . doi: 10.1017/S0305004100015681.

45. Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, *et al.* (2006): An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*. 31: 968–980.

46. Dale AM, Fischl B, Sereno MI (1999): Cortical surface-based analysis: I. Segmentation and surface reconstruction. *Neuroimage*. . doi: 10.1006/nimg.1998.0395.

47. Walhovd KB, Fjell AM, Giedd J, Dale AM, Brown TT (2016): Through Thick and Thin: a Need to Reconcile Contradictory Results on Trajectories in Human Cortical Development. *Cereb Cortex*. 1989: bhv301.

48. Zielinski BA, Prigge MBD, Nielsen JA, Froehlich AL, Abildskov TJ, Anderson JS, *et al.* (2014): Longitudinal changes in cortical thickness in autism and typical development. *Brain*. 137: 1799–1812.

49. Ducharme S, Albaugh MD, Nguyen T, Hudziak JJ, Mateos-pérez JM, Labbe A, *et al.* (2016): NeuroImage Trajectories of cortical thickness maturation in normal brain development — The importance of quality control procedures. *Neuroimage*. 125: 267–279.

50. Tamnes CK, Herting MM, Goddings A-L, Meuwese R, Blakemore S-J, Dahl RE, *et al.* (2017): Development of the Cerebral Cortex across Adolescence: A Multisample Study of Inter-Related Longitudinal Changes in Cortical Volume, Surface Area, and Thickness. *J Neurosci*. 37: 3402–3412.

51. Shaw P, Kabani NJ, Lerch JP, Eckstrand K, Lenroot R, Gogtay N, *et al.* (2008): Neurodevelopmental Trajectories of the Human Cerebral Cortex. *J Neurosci.* . doi: 10.1523/JNEUROSCI.5309-07.2008.

52. Fjell AM, Grydeland H, Krogsrud SK, Amlien I, Rohani DA, Ferschmann L (2015): Development and aging of cortical thickness correspond to genetic organization patterns. 112: 1–6.

53. A.M. F (2010): Structural brain changes in aging: Courses, causes and cognitive consequences. *Rev Neurosci*. 21: 187–221.

54. Mensen VT, Wierenga LM, van Dijk S, Rijks Y, Oranje B, Mandl RCW, Durston S (2016): Development of cortical thickness and surface area in autism spectrum disorder. *NeuroImage Clin*. 13: 215–222.

55. Thambisetty M, Wan J, Carass A, An Y, Prince JL, Resnick SM (2010): Longitudinal changes in cortical thickness associated with normal aging. *Neuroimage*. 52: 1215–1223.

56. Abrahams BS, Geschwind DH (2008): Advances in autism genetics: On the threshold of a new neurobiology. *Nat Rev Genet*. 9: 341–355.

57. Ecker C, Murphy D (2014): Neuroimaging in autism—from basic science to translational research. *Nat Rev Neurol*. 10: 82–91.

58. Geschwind DH, Levitt P (2007): Autism spectrum disorders: developmental disconnection syndromes. *Curr Opin Neurobiol*. 17: 103–111.

59. Marshall CR, Noor A, Vincent JB, Lionel AC, Feuk L, Skaug J, *et al.* (2008): Structural variation of chromosomes in autism spectrum disorder. *J Hum Genet*. 477–488.

60. Croen LA, Grether JK, Selvin S (2002): Descriptive Epidemiology of Autism in a California Population: Who Is at Risk? *J Autism Dev Disord*. 32: 217–224.

61. Seltzer MM, Shattuck P, Abbeduto L, Greenberg JS (2004): Trajectory of development in adolescents and adults with autism. *Ment Retard Dev Disabil Res Rev*. 10: 234–247.

62. Ronald A, Happé F, Bolton P, Butcher LM, Price TS, Wheelwright S, *et al.* (2006): Genetic heterogeneity between the three components of the autism spectrum: A twin study. *J Am Acad Child Adolesc Psychiatry*. 45: 691–699.

63. Hardan AY, Muddasani S, Vemulapalli M, Keshavan MS, Minshew NJ (2006): An MRI study of increased cortical thickness in autism. *Am J Psychiatry*. 163: 1290–1292.

64. Di Martino A, Yan CG, Li Q, Denio E, Castellanos FX, Alaerts K, *et al.* (2014): The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism. *Mol Psychiatry*. . doi: 10.1038/mp.2013.78.

65. Bethlehem RAI, Seidlitz J, Romero-Garcia R, Lombardo M V (2018): Using normative age modelling to isolate subsets of individuals with autism expressing highly age-atypical cortical thickness features. *bioRxiv*. . Retrieved from http://biorxiv.org/content/early/2018/01/23/252593.1.abstract.

66. Amaral DG, Schumann CM, Nordahl CW (2008): Neuroanatomy of autism. *Trends Neurosci*. 31: 137–145.

67. Jiao Y, Chen R, Ke X, Chu K, Lu Z, Herskovits E (2011): Predictive models of autism spectrum disorder based on brain regional cortical thickness. *Neuroimage*. 50: 589–599.

68. Ecker C, Andrews D, Dell'Acqua F, Daly E, Murphy C, Catani M, *et al.* (2016): Relationship between cortical gyrification, white matter connectivity, and autism spectrum disorder. *Cereb Cortex*. . doi: 10.1093/cercor/bhw098.

69. Ecker C, Ronan L, Feng Y, Daly E, Murphy C, Ginestet CE, *et al.* (2013): Intrinsic gray-matter connectivity of the brain in adults with autism spectrum disorder. *Proc Natl Acad Sci United States Am Sci*. 110: 13222–13227.

70. Moradi E, Khundrakpam B, Lewis JD, Evans AC, Tohka J (2017): Predicting symptom severity in autism spectrum disorder based on cortical thickness measures in agglomerative data. *Neuroimage*. 144: 128–141.

71. Doyle-Thomas KAR, Duerden EG, Taylor MJ, Lerch JP, Soorya L V., Wang AT, *et al.* (2013): Effects of age and symptomatology on cortical thickness in autism spectrum disorders. *Res Autism Spectr Disord*. 7: 141–150.

72. Giedd JN, Blumenthal J, Jeffries NO, Castellanos FX, Liu H, Zijdenbos A, *et al.* (1999): Brain development during childhood and adolescence: a longitudinal MRI study. *Nat Neurosci*. . doi: 10.1038/13158.

73. Hardan AY, Libove RA, Keshavan MS, Melhem NM, Minshew NJ (2009): A Preliminary Longitudinal Magnetic Resonance Imaging Study of Brain Volume and Cortical Thickness in Autism. *Biol Psychiatry*. 66: 320–326.

74. McAlonan GM, Cheung V, Cheung C, Suckling J, Lam GY, Tai KS, *et al.* (2005): Mapping the brain in autism. A voxel-based MRI study of volumetric differences and intercorrelations in autism. *Brain*. 128: 268–276.

75. McAlonan GM, Suckling J, Wong N, Cheung V, Lienenkaemper N, Cheung C, Chua SE (2008): Distinct patterns of grey matter abnormality in high-functioning autism and Asperger's syndrome. *J Child Psychol Psychiatry*. 49: 1287–1295.

Figure legends

Figure 1: Methodological overview. First, a normative model was estimated from cortical thickness derived from TD subjects (gray dots). Then we use this model to predict cortical thickness in ASD subjects (red dots). This allows us to estimate normative probability maps which show the regional deviations from the expected pattern in each subject. Finally, we generate a summary statistic quantifying the overall deviation for each subject by taking maximum deviation across brain using extreme value statistics. Abbreviations: CT = cortical thickness, NPM = normative probability map. See text for details.

Figure 2: Normative model of developmental changes of cortical thickness across the developmental range in the typical developing male cohort (the model was estimated using both genders). Cortical thickness was predicted using a trained normative model across the age range of six to thirty-one. The predicted cortical thickness map was thresholded so that only vertices that could accurately predict the true cortical thickness in the healthy cohort under cross-validation were retained (Pearson correlation, $p < 0.05$, FDR). Blue vertices and yellow indicate reduced and increased CT respectively. Moreover, the predicted cross-sectional developmental trajectories of CT in four randomly-selected vertices are shown.

Figure 3: Vertex-wise group differences between ASD and TD cohorts after FDR correction ($p < 0.05$). The green circles indicate the regions show the vertex-wise group difference. No vertices survived after FDR correction in vertex-wise group differences map between ASD and TD in female and male separately.

Figure 4: Overlap of vertex-wise negative deviation across each cohort and schedule. This map shows the spatial distribution of individual subjects with significant deviations in each vertex after FDR correction. The proportion of subjects contributing to each map is also shown (i.e. the proportion of subjects having deviations surviving FDR correction).

Figure 5: Overlap of vertex wise positive deviation across each cohort and schedule. See caption to Figure 4 for further details

Figure 6: Extreme value histogram and distribution

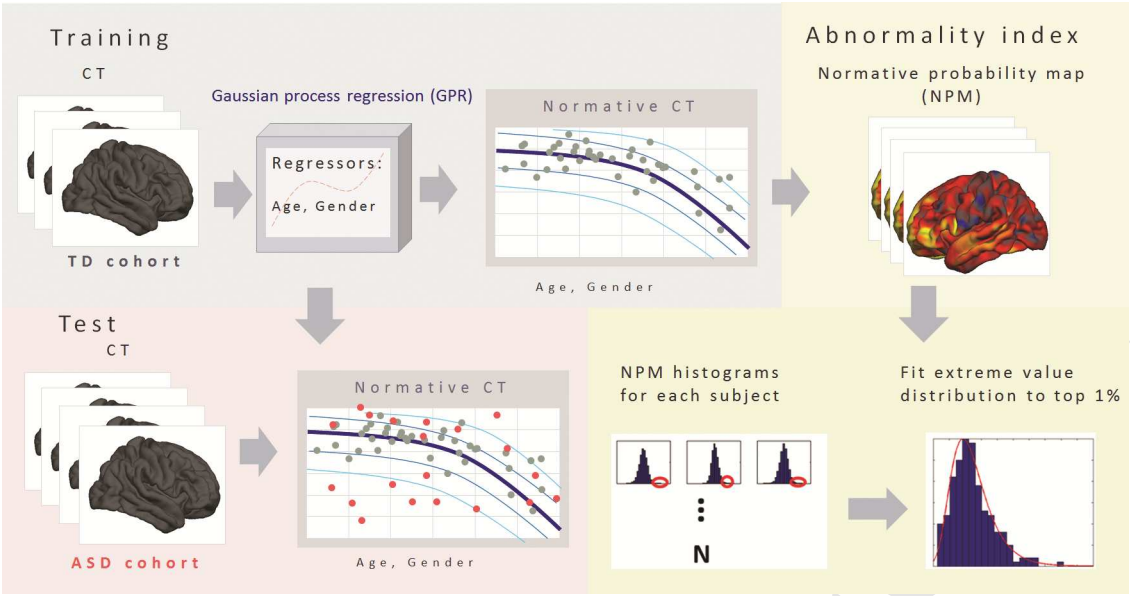
Figure 7: Regional extreme value deviation correlation with ASD symptoms-female ($p < 0.05$) according to the Desikan-Killany parcellation scheme. Blue and yellow regions indicate negative and positive association with ASD symptoms, respectively. Green circles indicate the regions that survived after FDR correction.

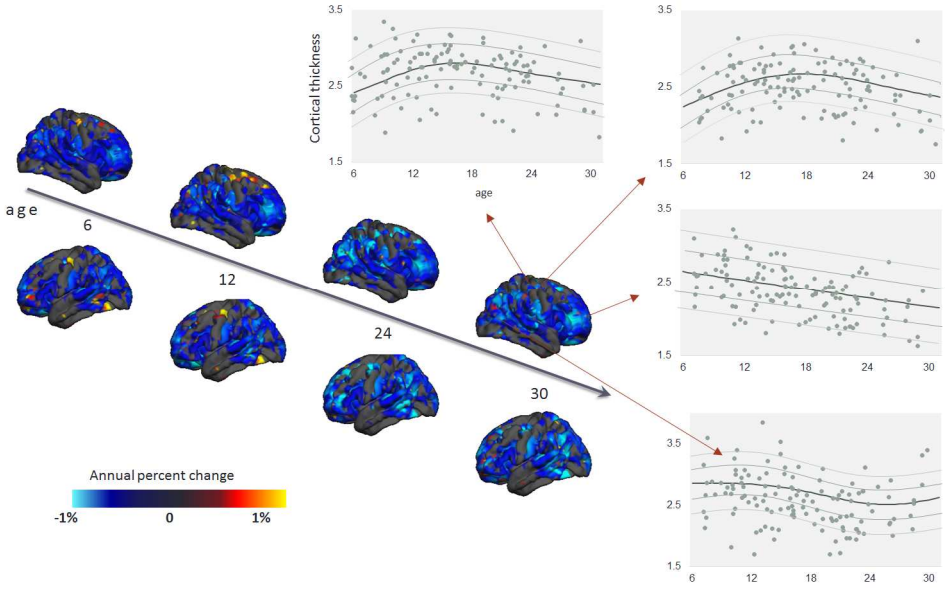
Figure 8: Regional extreme value deviation correlation with ASD symptom for males ($p < 0.05$) according to the Desikan-KIllany parcellation scheme

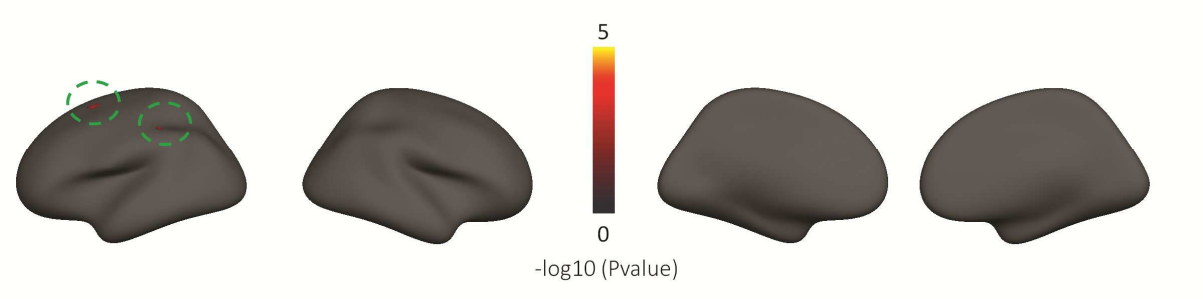
Tables

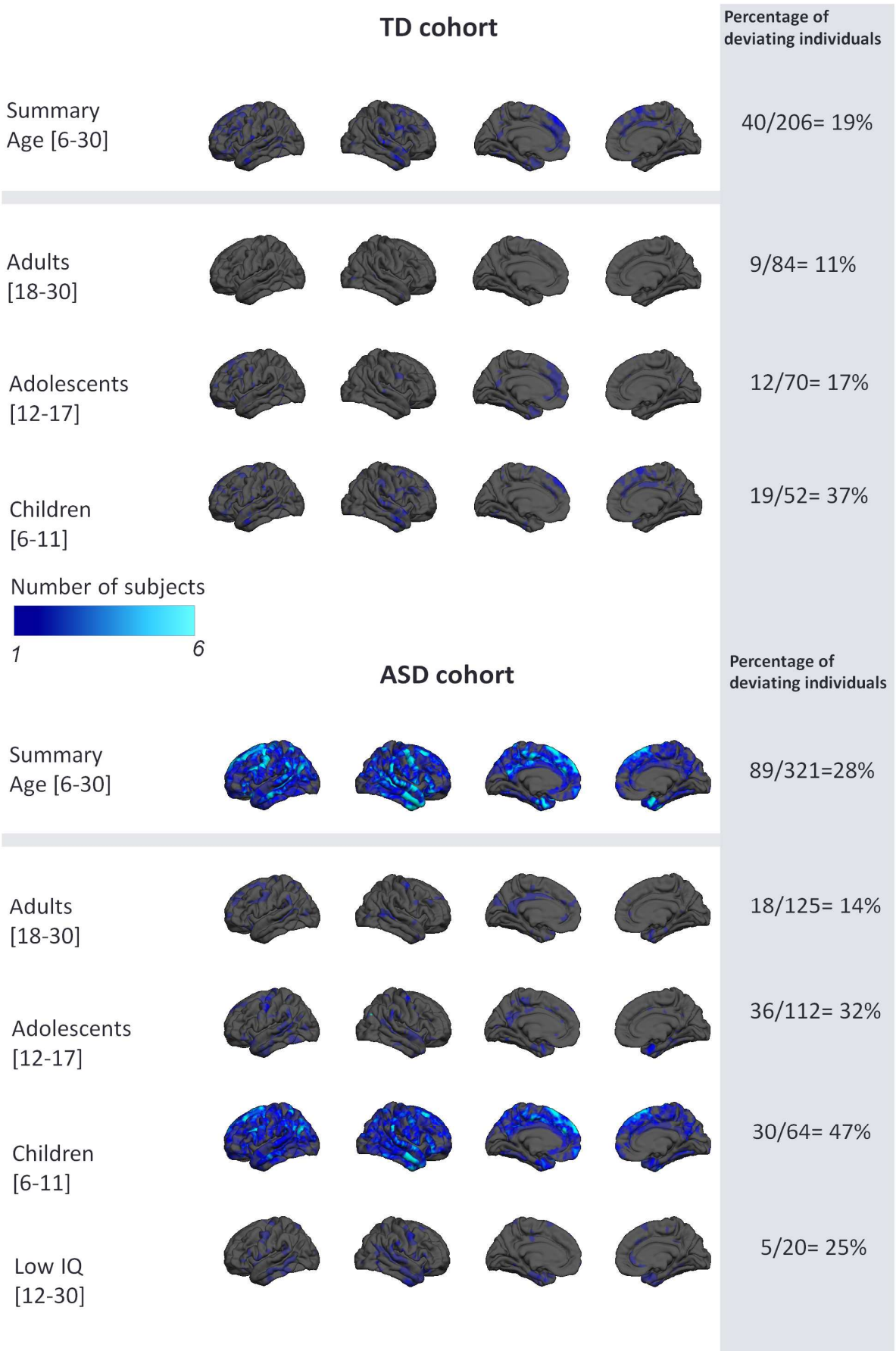
Table 1: Clinical characteristics

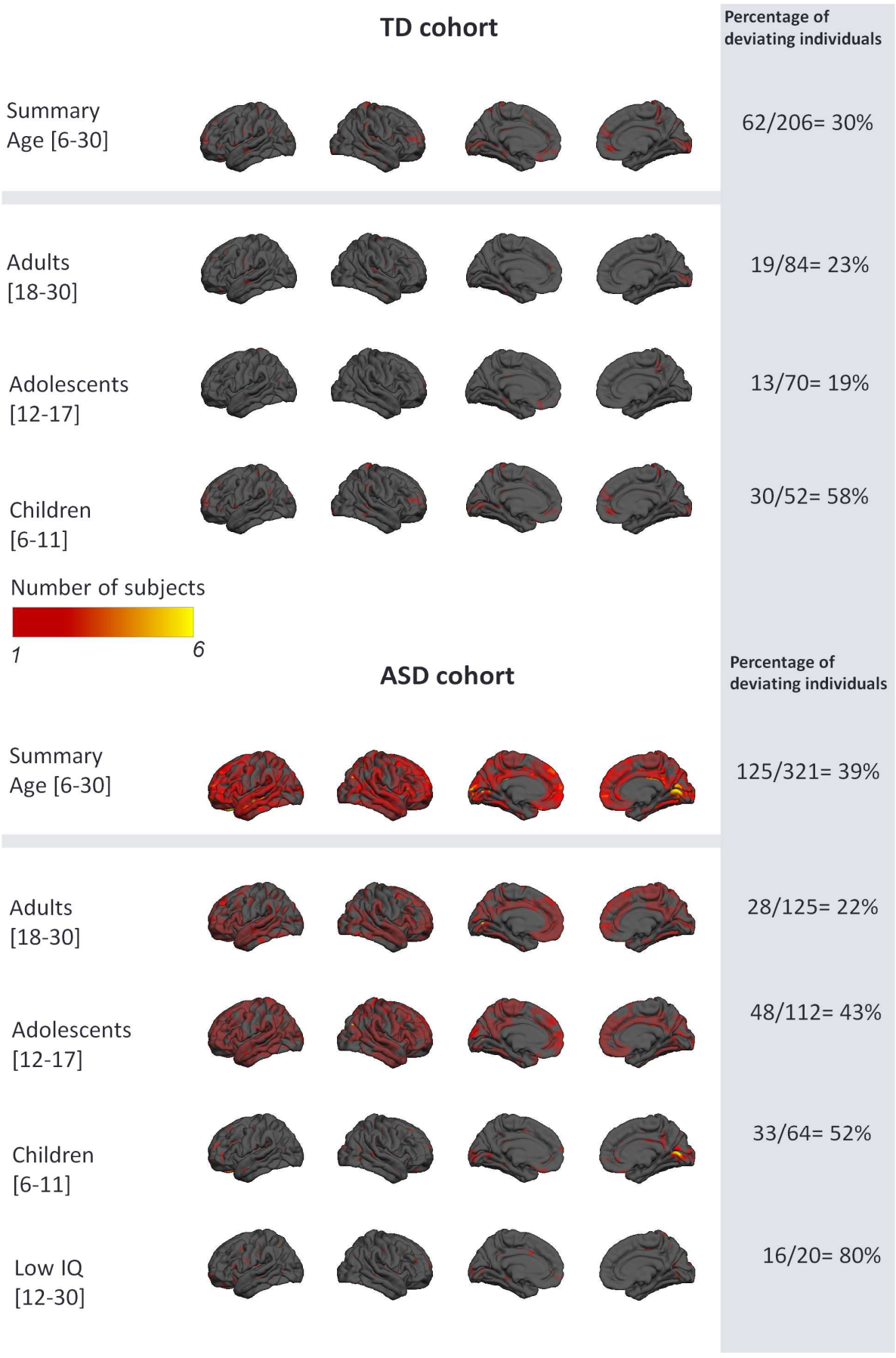
variable	***ASD*** *n=321 (89 female)*	***TD*** *n=206 (79 female)*	***P value***
Age, mean, [SD]	17.01 [5.79]	17.14 [5.97]	0.93 (ns)
IQ , mean [SD]			
Global IQ	100.89 [18.53] N=316	108.22 [14.24]	0.00
Performance IQ	101.65 [20.14] N=316	108.26 [15.72]	0.00
Verbal IQ	99.64 [18.53] N=313	107.32 [16.13]	0.00
ADI-R [SD]	N=308		
Social	16.20 [6.71]		
Communication	13.11 [5.69]		
Repetitive Behavior	4.32 [2.69]		
ADOS [SD]	N=258		
Total	5.12 [2.77]		
Social	5.78 [2.62]		
Repetitive Behavior	4.78 [2.76]		
Schedule			
A: Adults	125	84	
B: Adolescence	112	70	
C. Children	64	52	
D. IQ < 70	20	-	

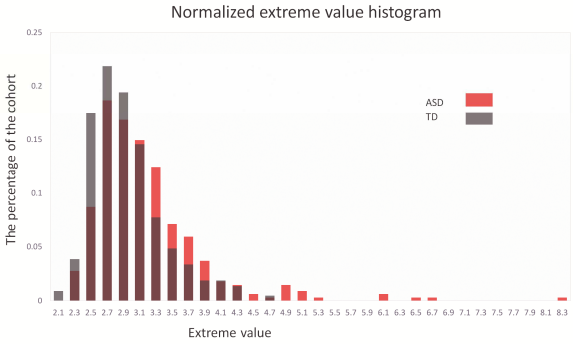


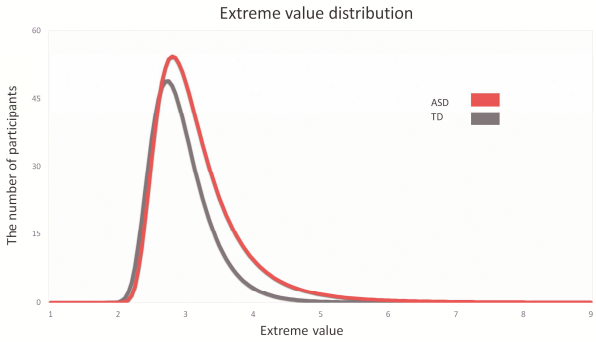


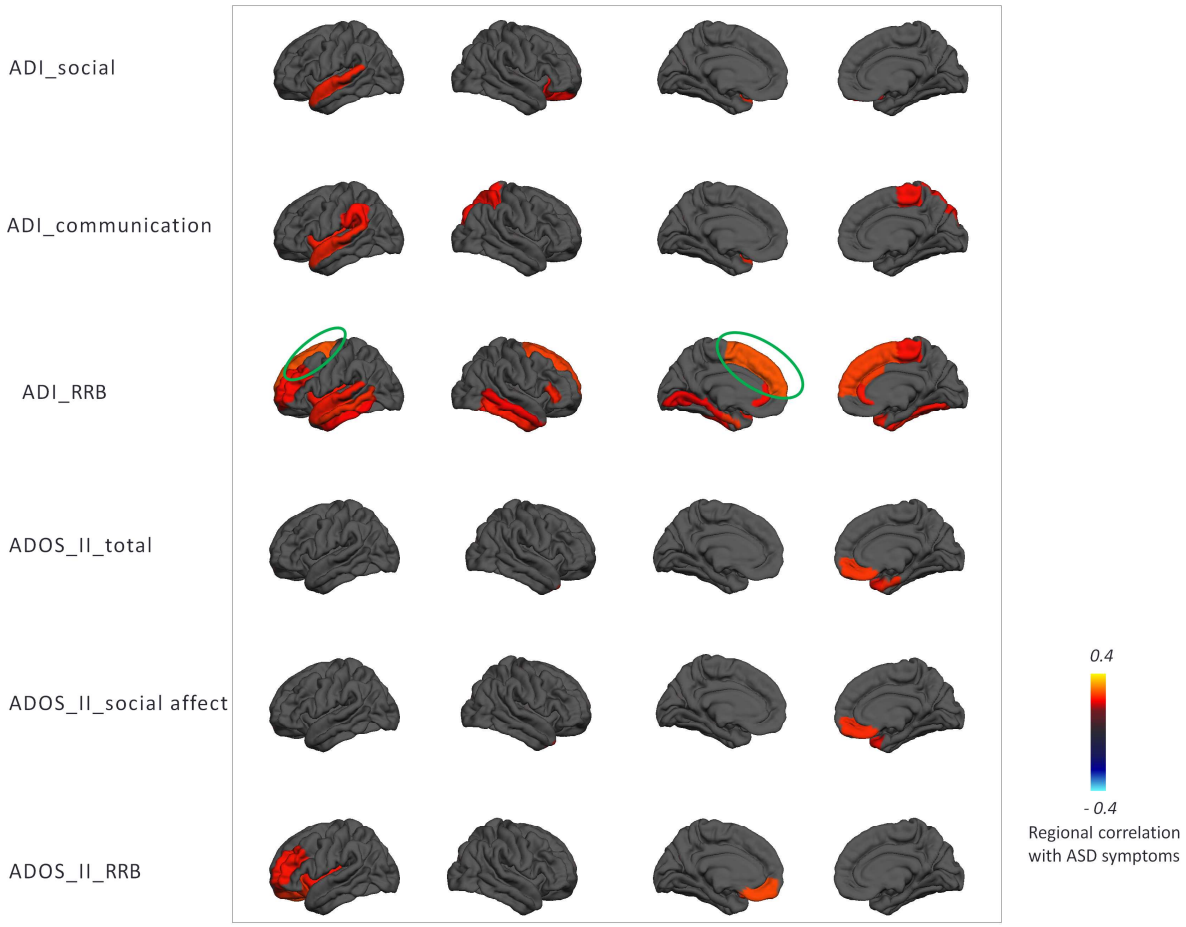


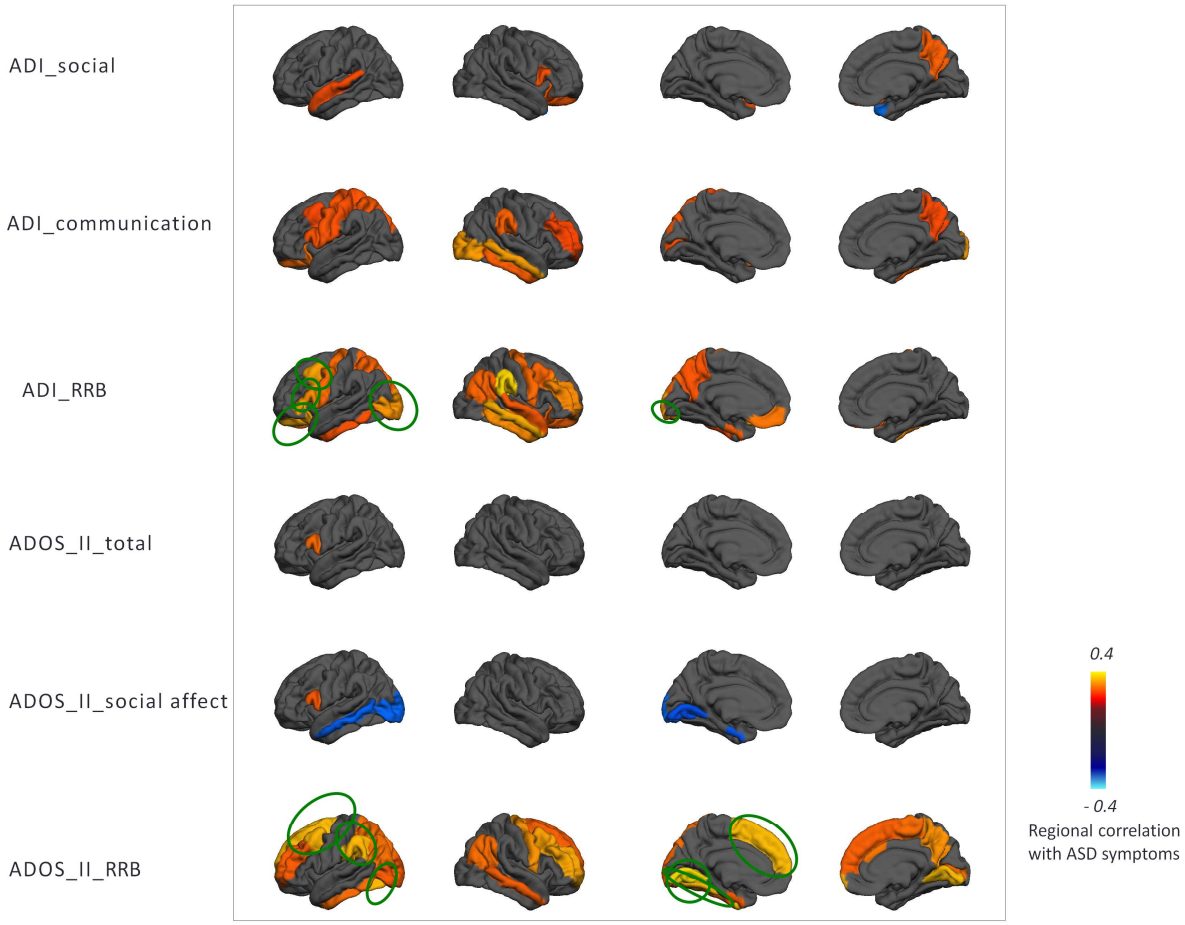

TD cohort
Percentage of deviating individuals
Summary Age [6-30]
40/206= 19%
Adults [18-30]
9/84= 11%
Adolescents [12-17]
12/70= 17%
Children [6-11]
19/52= 37%
Number of subjects
1
6
ASD cohort
Percentage of deviating individuals
Summary Age [6-30]
89/321=28%
Adults [18-30]
18/125= 14%
Adolescents [12-17]
36/112= 32%
Children [6-11]
30/64= 47%
Low IQ [12-30]
5/20= 25%


TD cohort
Percentage of deviating individuals
Summary Age [6-30]
62/206= 30%
Adults [18-30]
19/84= 23%
Adolescents [12-17]
13/70= 19%
Children [6-11]
30/52= 58%
Number of subjects
1
6
ASD cohort
Percentage of deviating individuals
Summary Age [6-30]
125/321= 39%
Adults [18-30]
28/125= 22%
Adolescents [12-17]
48/112= 43%
Children [6-11]
33/64= 52%
Low IQ [12-30]
16/20= 80%









Dissecting the Heterogeneous Cortical Anatomy of Autism Spectrum Disorder Using Normative Models

Supplementary Information

Supplementary methods

Exclusion criteria

Exclusion criteria for the Longitudinal European Autism Project (LEAP) study included a history of substance abuse and standard neuroimaging contraindications (e.g. claustrophobia, metal implants). Individuals were also excluded if they had a history of bipolar disorder or psychosis. In contrast to case-control studies that aim to detect consistent group differences, here we were interested in characterizing the heterogeneity within autism spectrum disorder (ASD) at the level of the individual. Therefore, we did not exclude other comorbidities in the clinical group because up to 70% of ASD individuals have one or more psychiatric conditions (1) and 30-50% of individuals with ASD are on stable medications (2).

Magnetic Resonance Imaging

A high resolution T1-weighted image was acquired from each participant with a standard Alzheimer's Disease Neuroimaging Initiative (ADNI) sequence (3), matched across scanning sites. Cortical thickness was estimated from the high-resolution T1-weighted image for each subject using FreeSurfer version 5.3 (http://surfer.nmr.mgh.harvard.edu/). Prior to analysis, all surface reconstructions were visually assessed for reconstruction errors by at least 3 independent raters. We excluded a small number of scans with severe artifacts (e.g. caused by head motion). The rest of the scans were included 'as is', i.e. we did not allow manual edits to reduce the possibility of bias (e.g. due to individual differences in operator skill). Cortical thickness maps were then smoothed with a 10-mm surface-based Gaussian kernel.

Gaussian process regression

As mentioned in the main text, Gaussian process regression (GPR) (4) was used to estimate separate normative models of cortical thickness (CT) and surface area (SA) at each vertex on the cortical surface. Whilst other methods are also suited to this purpose (e.g. Bayesian polynomial regression), in preliminary testing we found that GPR provides superior estimation of the mean and the ability to map the variation across the cohort through centiles of predictive confidence. We refer the reader elsewhere for a full treatment of Gaussian processes (5, 6) but briefly, a Gaussian process (GP) specifies a distribution over

functions, such that any finite number of elements has a joint Gaussian distribution. They are excellent tools for Bayesian regression: given a dataset specified by $\mathcal{D} = \{\mathbf{x}_i, y_i \in\}_{i=1}^{N}$ – where $\mathbf{x}_i$ are D-dimensional vectors of covariates, N is the total sample size and $y_i \in \mathbb{R}$ are response variables – the response variables are predicted using a potentially nonlinear regression model with additive Gaussian noise, i.e.: $y_i = f_i + \epsilon_i$ where $\epsilon_i \sim N(0, \sigma_n^2)$. Inference then proceeds by placing a GP prior over this function then computing the posterior distribution using the canonical GPR predictive equations (5). This prior is uniquely specified by a mean ($m(x)$) and covariance ($k(x, x')$) function. Here, without loss of generality we choose a mean function equal to zero and a generic covariance function combining linear and non-linear terms, i.e.:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T \mathbf{x}_j + \sigma_f \exp\left(-\frac{1}{2}(\mathbf{x}_i - \mathbf{x}_j)^T \mathbf{\Lambda} (\mathbf{x}_i - \mathbf{x}_j)\right)$$

Where σ_f is a signal amplitude parameter for the nonlinear component and $\mathbf{\Lambda}$ is a diagonal matrix with ℓ_d^{-2} along the leading diagonal. These are 'automatic relevance determination' parameters (5) that can down-weight irrelevant dimensions in the input space or emphasize important dimensions. Training a GP model refers to finding the optimal values for the model parameters which are: $\ell_1, \ldots, \ell_D, \sigma_n$ and σ_f. This is conveniently achieved by maximizing the logarithm of the model evidence (i.e. the denominator of Bayes rule). Finally, we compute a single subject Z-statistic image for each subject (i) and at each brain location (j) by computing:

$$z_{ij} = \frac{y_{ij} - \hat{y}_{ij}}{\sqrt{\sigma_{ij}^2 + \sigma_{nj}^2}}$$

Here, $\hat{y}_{ij}$ is the predicted mean and predicted variance, σ_{ij}, which is combined with the true response (y_{ij}) and variance learned from the TD distribution (σ_{nj}). Because we estimate a separate noise parameter for each vertex, this should accommodate regional differences in population variation (for example, the estimated variance parameter will be higher in the regions where there is greater variation across individuals).

Cross-validation

To assess generalization, we used 10-fold cross-validation where we partitioned the data into 10 'folds' and repeatedly trained the model on 90% of the data, withholding the remaining 10% for estimating generalization performance. This was repeated 10 times so that each partition was excluded once. This

procedure is standard in machine learning and is known to provide approximately unbiased estimates of the true generalization ability.

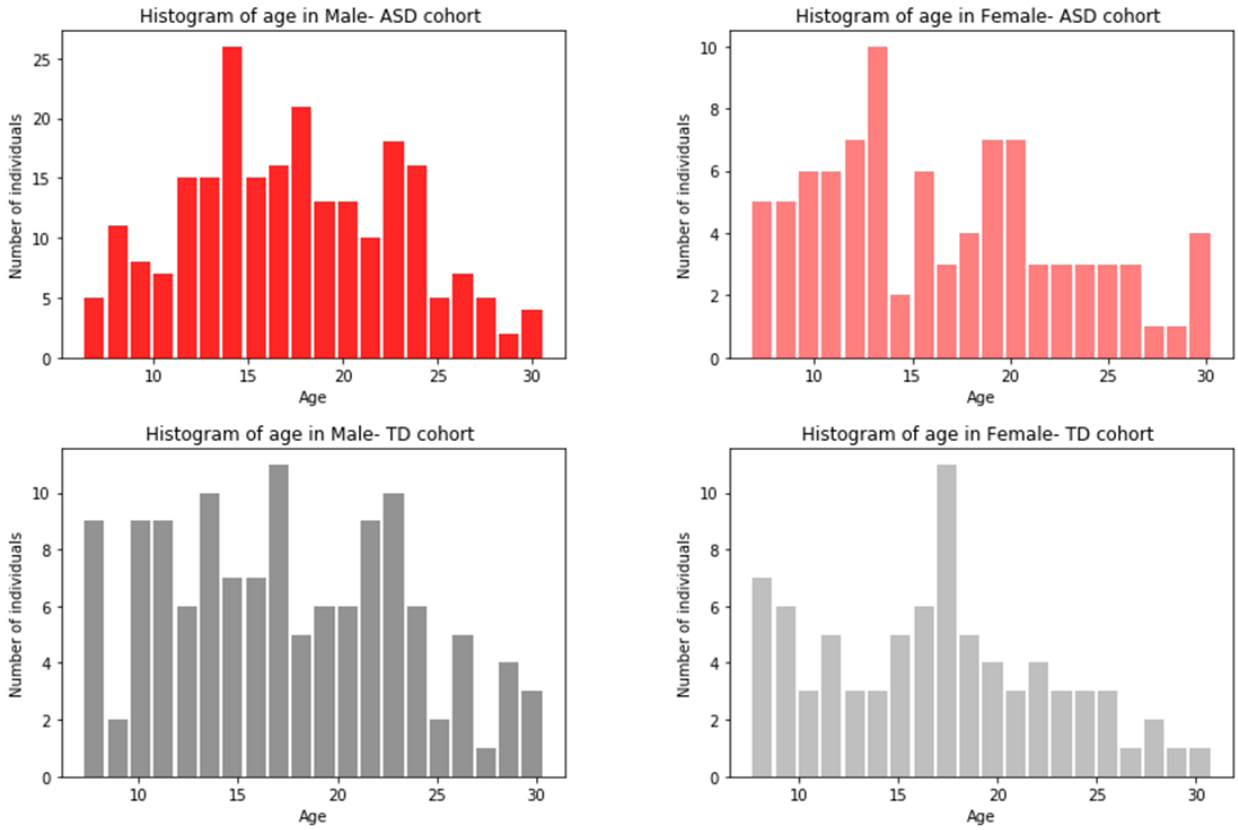
Post-hoc investigation of potential confounding variables

To investigate the potential confounding effect of various potential confounding variables, we performed several tests. As described in the main text, we first estimated a normative model for CT additionally including scanning site, IQ and the FreeSurfer Euler number (EN) (7) as covariates. In addition, we additionally performed several post-hoc tests for potential confounding variables including scan quality, IQ and comorbid attention deficit/hyperactivity disorder (ADHD) symptoms. However, we emphasize strongly that these should be considered as illustrative only, because our study design does not allow us to determine the direction of cause-effect relationships. For example, it is reasonable to expect that subjects that show the most atypical cortical anatomy may also express the highest level of symptoms, have the most intellectual impairment and be the most likely to suffer from comorbid symptoms. First, to assess the possibility of scan quality (e.g. due to excessive head motion in the scanner) influencing our results, we correlated (using Spearman correlation) the deviations from the model with the Euler number. The EN summarizes the topological complexity of the estimated cortical surface and has been proposed as a proxy measure of scan quality (8) . However, this is an indirect measure in that it does not model scan quality directly, it should be considered with caution since many other variables can potentially influence EN, including age, atypicalities in cortical anatomy and disorder severity. Therefore, we also correlated EN with age and with ASD symptoms. In addition, we correlated the deviations from the normative with measures of full-scale IQ (see (9)) and with measures of comorbid ADHD symptoms derived from the Development and Well-being Assessment (10). We refer the reader elsewhere for a detailed description of these measures (3, 9).

Supplementary results

Age histogram

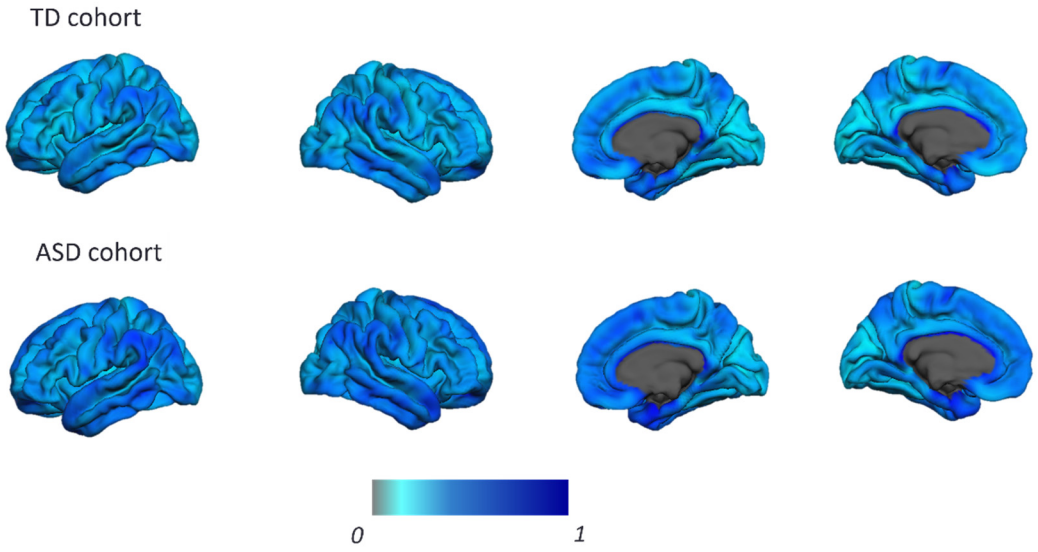
Figure S1 shows the distribution of subject ages across diagnoses.



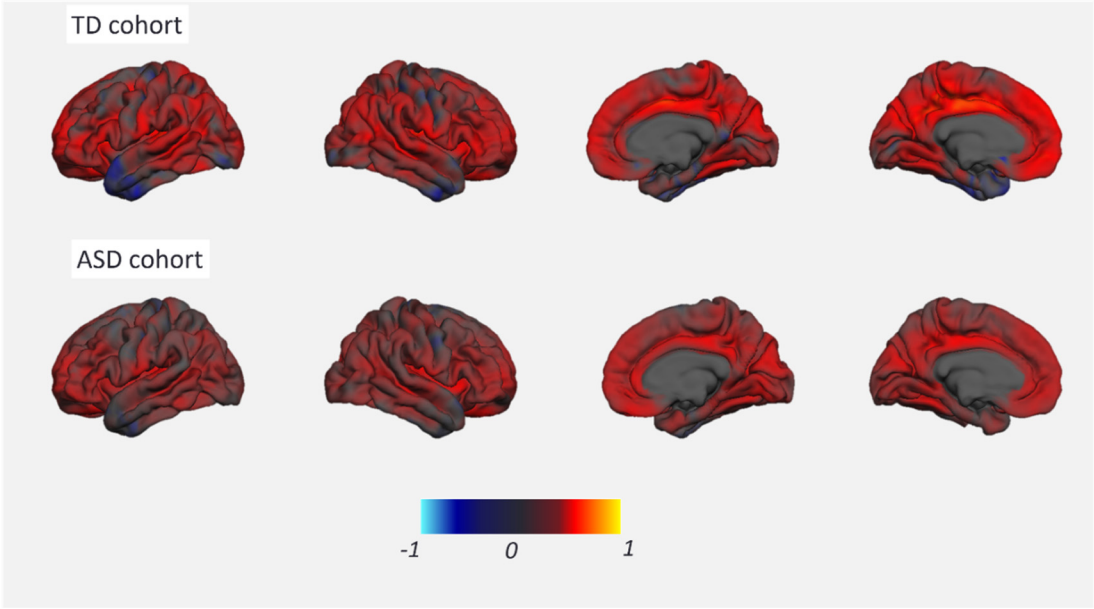
Supplementary Figure S1: Histogram of age of Female and Male individuals across TD and ASD cohort.

Model fit evaluation

Figure S2 shows the mean accuracy of the normative model for predicting CT in typically developing (TD) and ASD participants, both in terms of root mean squared error (Figure S2A) and correlation between true and predicted CT values (Figure S2B).



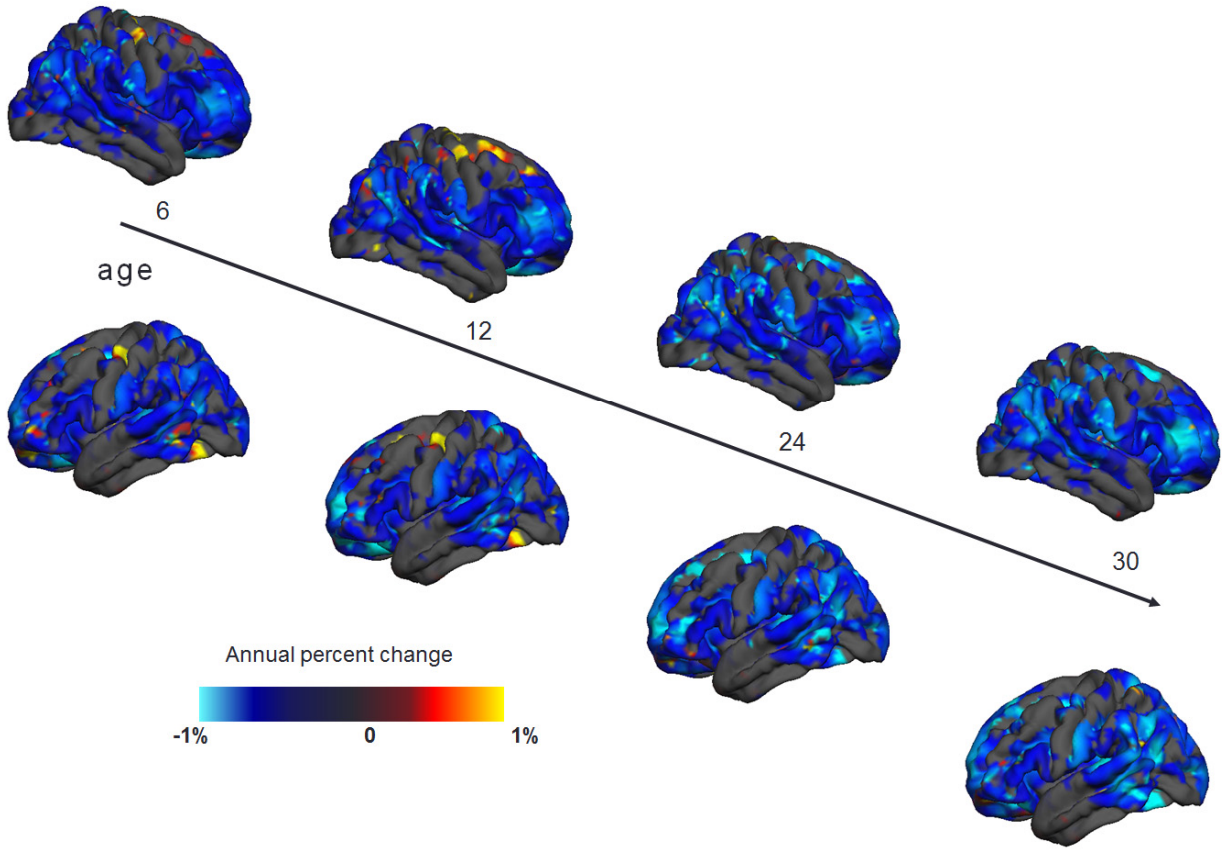
Supplementary Figure S2.A: Root mean square error of true and prective mean of cortical thickness in TD cohort.



Supplementary Figure S2.B: The correlation between true and prective mean of cortical thickness in TD and ASD cohort.

Normative developmental changes for cortical thickness in females

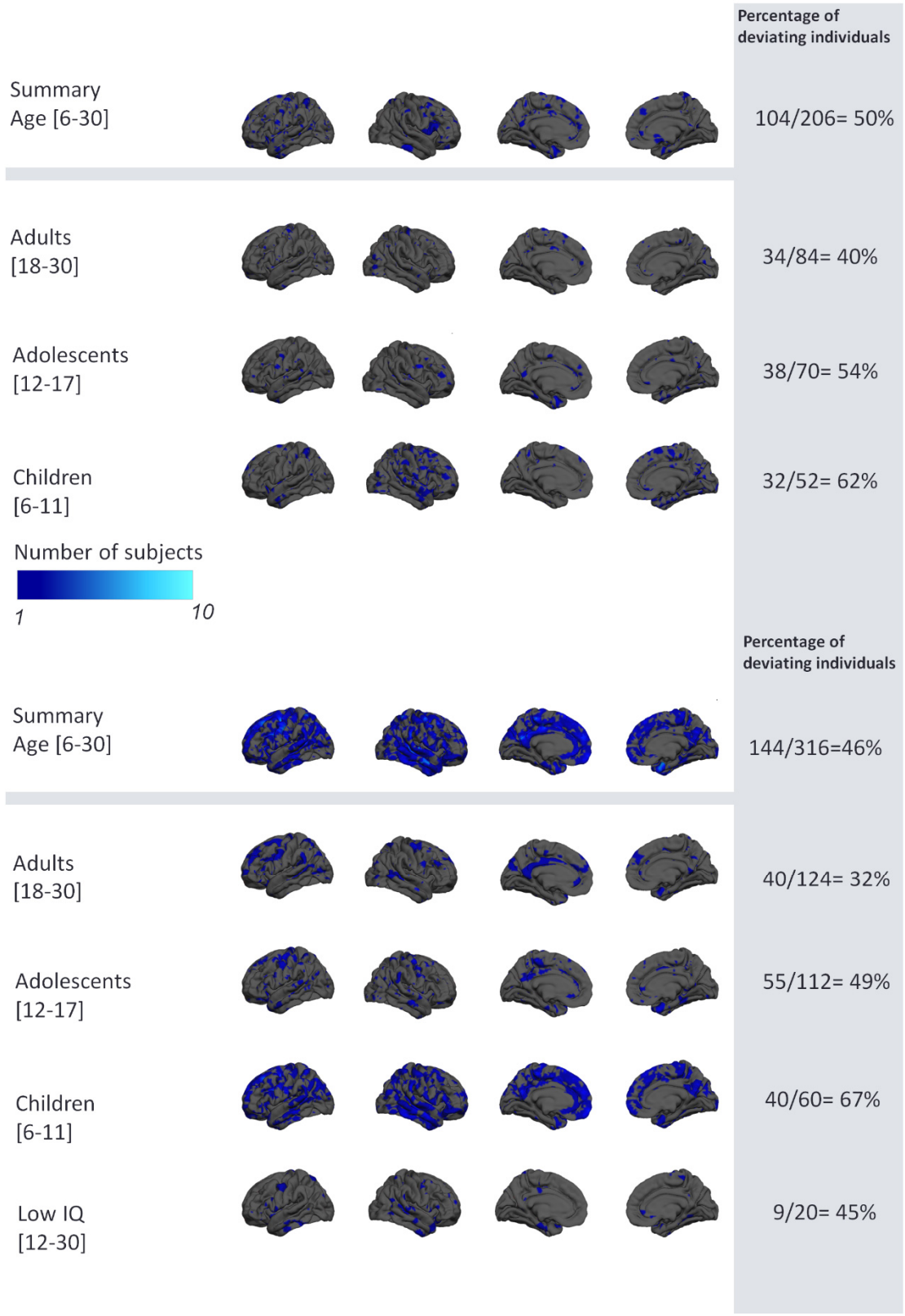
Supplementary Figure S3 shows the predictions made by the normative model for changes in female TD subjects (see Figure 2 in the main text for males).



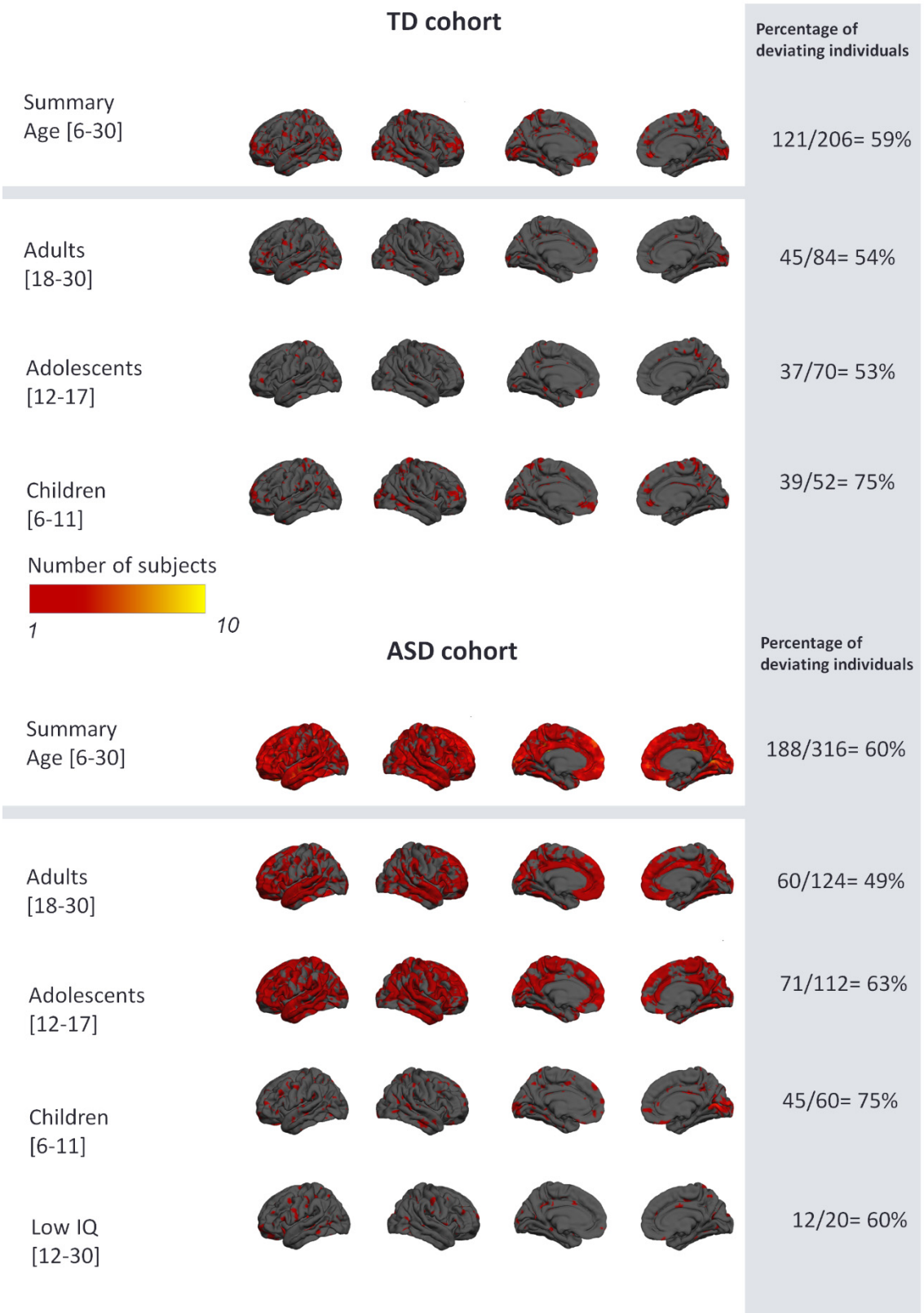
Supplementary Figure S3: Normative model of developmental changes of cortical thickness across the developmental range in the typical developing female cohort. Cortical thickness was predicted using a trained normative model across the age range of six to thirty-one. The predicted cortical thickness map was thresholded so that only vertices that could accurately predict the true cortical thickness in the healthy cohort under cross-validation were retained (Pearson correlation, $p < 0.05$, FDR). Blue vertices and yellow indicate reduced and increased CT respectively.

Normative model of cortical thickness using age, gender, site, IQ and FreeSurfer Euler number as covariates

Supplementary Figures S4 and S5 show deviations from the normative model for CT re-estimated after additionally including IQ addition and scanning site dummy variables and FreeSurfer Euler number as covariates, separately for positive (Figure S4) and negative (Figure S5) deviations. The differences between this model and the original model are negligible and all the conclusions remain unchanged.



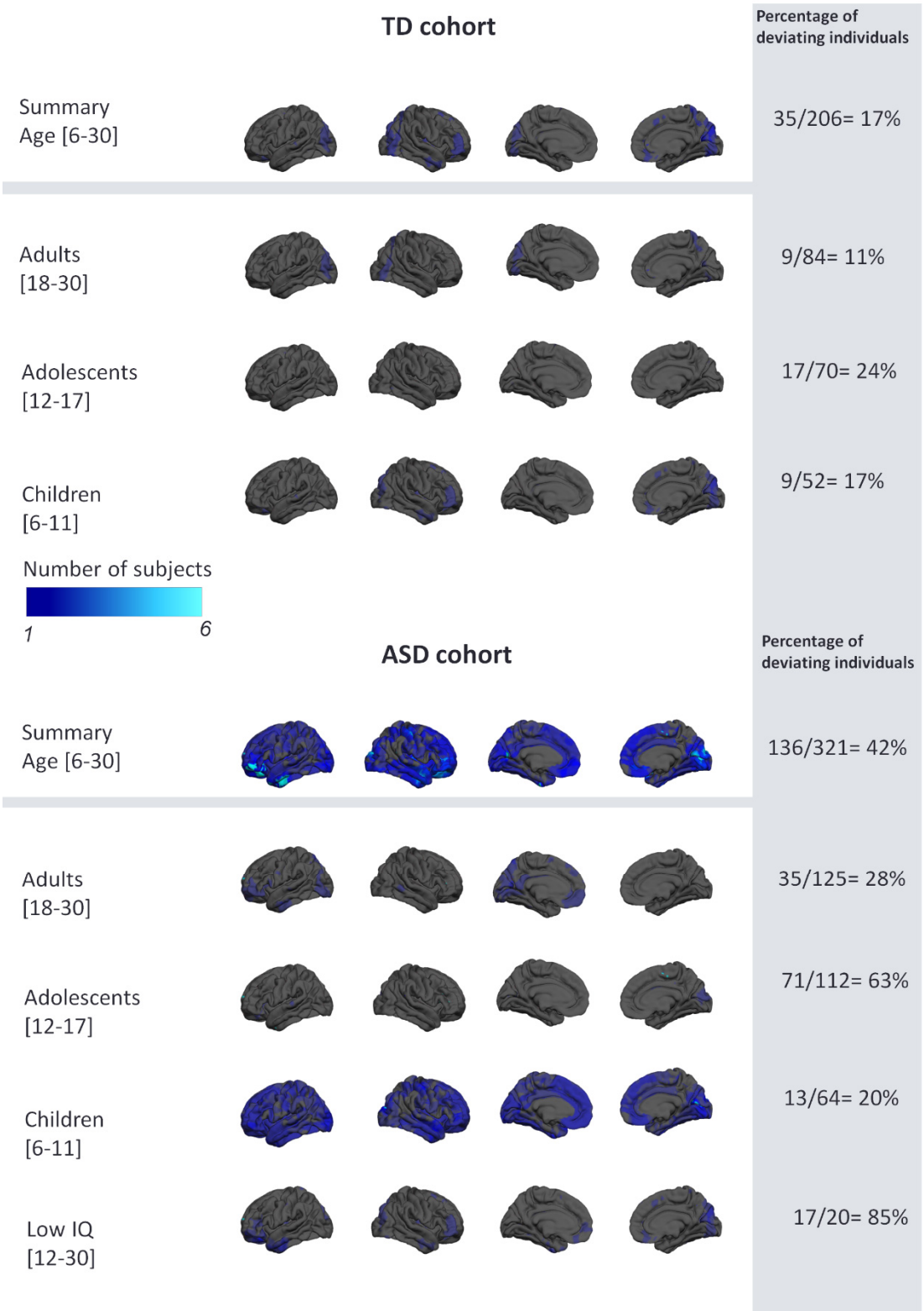
Supplementary Figure S4: Overlap of vertex-wise negative deviation across each cohort and schedule. This map shows the number of subjects with significant deviations in each vertex after FDR correction.



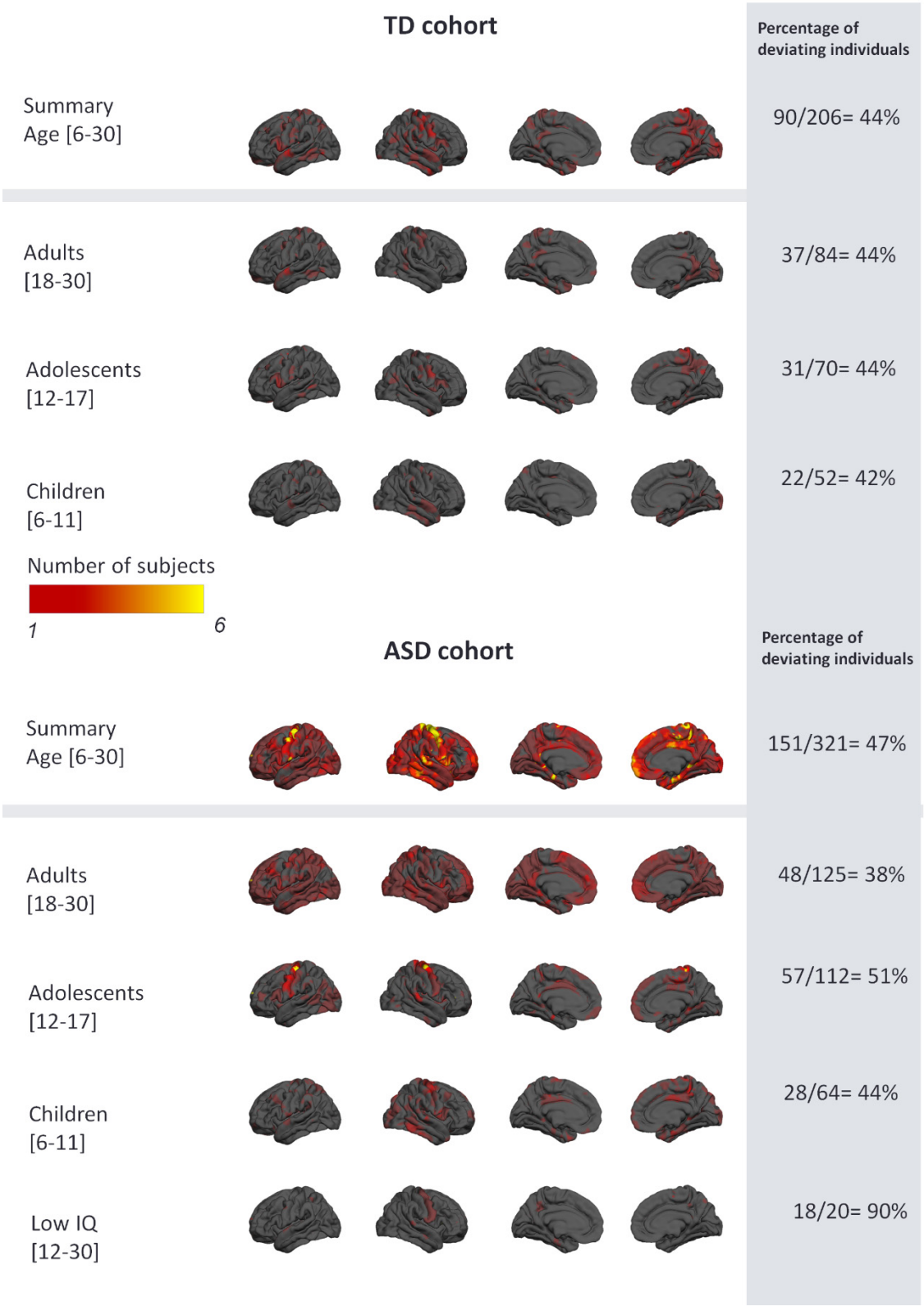
Supplementary Figure S5: Overlap of vertex wise positive deviation across each cohort and schedule.

Normative model of surface area

Supplementary Figure S6 and Figure S7 show deviations from the normative model for CT estimated using age and gender as covariates, separately for positive (Figure S6) and negative (Figure S7) deviations. The overlap of deviating voxels shows a similar but slightly different pattern relative to CT.



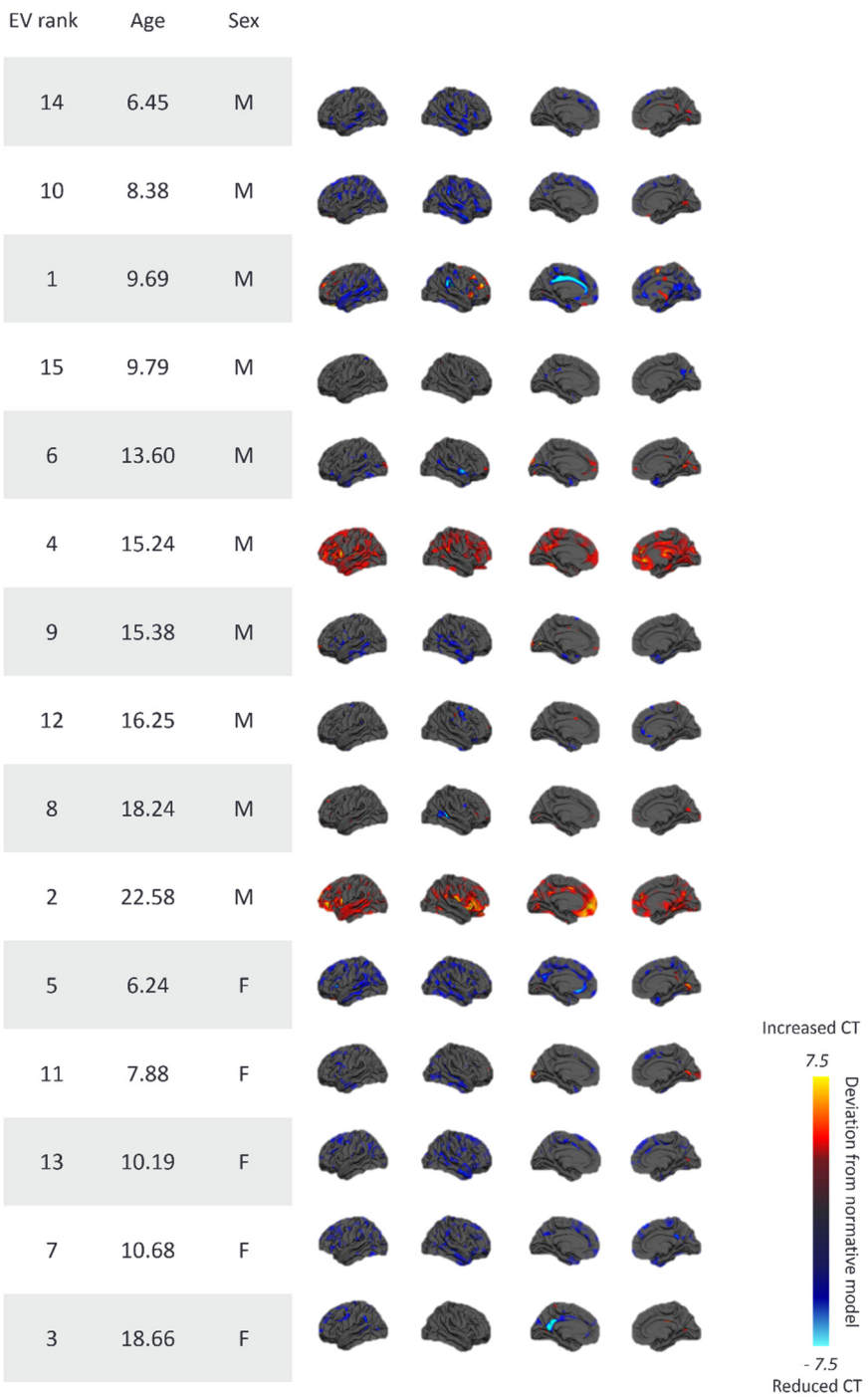
Supplementary Figure S6: Overlap of vertex-wise negative deviation in surface area normative model across each cohort and schedule. This map shows the number of subjects with significant deviations in each vertex after FDR correction.



Supplementary Figure S7: Overlap of vertex wise positive deviation across each cohort and schedule.

Individual subject deviations

Supplementary Figure S8 shows the top 15 subjects deviating from the normative pattern for CT.



Supplementary Figure S8: NMPs of top fifteen deviating individuals from normative CT model. These subjects who belong to ASD cohort have highly individualized patterns of deviation with respect to brain regions and different sign of the deviation.

Individual subject deviations

Supplementary Table S1 shows the clinical characteristics of the sample separately for each site.

Supplementary Table S1: Clinical characteristics for each site.

	Cambridge		KCL		Mannheim		Nijmegen		Rome		Utrecht	
Variable	***ASD***	***TD***	***ASD***	***TD***	***ASD***	***TD***	***ASD***	***TD***	***ASD***	***TD***	***ASD***	***TD***
Age, mean, [SD]	17.6 [5.8]	17.0 [6.4]	16.9 [5.9]	18.7 [6.6]	15.4 [3.4]	15.3 [3.6]	16.2 [5.6]	15.0 [4.1]	24.9 [2.9]	24.9 [3.5]	16.8 [5.5]	16.7 [6.2]
IQ, mean [SD]												
Global IQ	106 [19]	114 [11]	100 [21]	111 [16]	102 [13]	109 [15]	97 [18]	101 [14]	101 [14]	106 [10]	105 [13]	111 [8]
Performance IQ	109 [21]	116 [12]	100 [20]	110 [16]	104 [15]	113 [13]	97 [22]	102 [18]	104 [18]	103 [15]	107 [17]	109 [12]
Verbal IQ	103 [17]	109 [11]	99 [20]	111 [18]	101 [15]	101 [14]	97 [19]	100 [15]	98 [16]	109 [8]	105 [14]	113 [14]
ADI-R [SD]												
Social	17.2 [6.6]	-	17.9 [6.4]		15.2 [7.2]	-	14.4 [6.4]	-	11.4 [5.7]	-	16.3 [5.8]	-
Communication	14.5 [2.7]	-	15.2 [5.4]	-	10.4 [4.9]	-	12.7 [5.4]	-	9.3 [5.4]	-	11.3 [5.4]	-
Repetitive Behavior	5.0 [2.7]	-	5.0 [2.4]	-	5.1 [3.8]	-	2.9 [2.1]	-	5.3 [2.3]	-	3.5 [2.7]	-
ADOS [SD]												
Total	5.2 [2.4]	-	5.1 [2.9]	-	-	-	5.3 [2.7]	-	-	-	4.8 [2.7]	-
Social	6.3 [1.8]	-	5.3 [2.8]	-	-	-	6.1 [2.5]	-	-	-	5.5 [2.7]	-
Repetitive Behavior	4.5[2.5]	-	5.7 [2.6]	-	-	-	3.7[2.7]	-	-	-	4.3[2.4]	-
Schedule												
A: Adults	16	9	50	33	5	5	23	10	17	10	14	10
B: Adolescents	15	6	34	18	20	11	31	27	0	0	12	8
C. Children	7	7	26	9	3	5	20	18	0	0	8	13
D. IQ < 70	2	-	12	-	0	-	6	-	0	-	0	-

Scan quality associations with extreme deviations and age

In Supplementary Table S2, we show the correlation of deviations with the FreeSurfer Euler number. Note that smaller EN values are indirectly associated with a lower scan quality.

Supplementary Table S2: Correlation between extreme deviations and Euler number.

	Euler number in ASD cohort	Euler number in TD cohort	Euler number overall
Extreme deviations	-0.57 *	-0.62 *	-0.58 *
Age	0.39 *	0.38 *	0.38 *
ADI Social	-0.18 *	-	-
ADI Communication	-0.25 *	-	-
ADI RRB	-0.25 *	-	-
ADOS Total	-0.11	-	-
ADOS Social	-0.02	-	-
ADOS RRB	-0.21 *	-	-

This shows that EN was correlated with the deviations in both ASD and TD cohorts, all ADI symptom domains, and ADOS repetitive behaviors. Regarding alternative potentially confounding variables, the extreme value deviations were weakly negatively correlated with IQ in the ASD cohort (ρ = -0.16, $p<0.05$), but not in the TD cohort (ρ = 0.08, n/s). They were also correlated with some ADHD symptom scales (Supplementary Table S3).

Supplementary Table S3: Correlation between extreme deviations and ADHD score.

	ADHD Inattentive parent in ASD	ADHD Inattentive parent in TD	ADHD Inattentive
Extreme deviations	0.20 *	-0.01	0.22 *
	ADHD Hyperimpulsive parent in ASD	**ADHD Hyperimpulsive parent in TD**	**ADHD Hyperimpulsive**
Extreme deviations	0.24 *	0.16	0.26 *

Taken together these results reinforce the cautions noted above, and preclude a definitive assessment of the degree to which any one particular confounding variable may have influenced our results.

Supplementary Table S4 shows the clinical characteristics from the top deviating subjects from the normative model for CT, all of whom have ASD

Supplementary Table S4: Clinical characteristics of the top fifteen deviating participants from normative CT model. These subjects who belong to ASD cohort, have highly individualized patterns of deviation with respect to brain regions and different sign of the deviation.

		Schedule	Site	Sex	VIQ	PIQ	FSIQ	Age	ADI-social	ADI-communication	ADI-RRB	ADOS-TOTAL	ADOS-SA	ADOS-RRB
1	ASD	Children	Mannheim	M	-	69	-	9.69	20	10	8	-	-	-
2	ASD	Adults	Rome	M	70	79	76	22.58	18	11	3	-	-	-
3	ASD	Adults	Nijmegen	F	79	66	74	18.66	4	6	0	3	4	6
4	ASD	Adolescent	Utrecht	M	86	106	96	15.24	10	9	2	5	6	6
5	ASD	Children	KCL	F	90	109	100	6.24	17	17	7	10	10	8
6	ASD	Adolescent	Nijmegen	M	114	124	118	13.6	15	15	2	-	-	-
7	ASD	Children	KCL	F	116	133	127	10.68	10	9	8	1	2	6
8	ASD	Adults	KCL	M	87	86	84	18.42	-	-	-	2	2	6
9	ASD	IQ < 70	KCL	M	58	59	56	15.83	22	21	9	10	10	10
10	ASD	Children	KCL	M	111	106	110	8.38	16	18	5	6	3	9
11	ASD	Children	Utrecht	F	93	98	95	7.88	23	17	3	7	5	9
12	ASD	IQ < 70	Nijmegen	M	52	50	54	16.25	27	21	6	10	10	8
13	ASD	Children	KCL	F	-	-	-	10.19	0	8	2	6	7	6
14	ASD	Children	KCL	M	99	107	104	6.45	14	12	5	2	3	1
15	ASD	Children	Cambridge	'M	106	128	119	9.79	16	15	2	4	5	6

Supplementary Table S5 shows the clinical characteristics from the top deviating subjects from the normative model for SA.

Supplementary Table S5: Clinical characteristics of the top fifteen deviating participants from normative surface area model. While 80% of the individuals belong to ASD cohort with highly heterogeneous profiles, there are several individuals in the list who belong to TD cohort.

		Schedule	Site	Sex	VIQ	PIQ	FSIQ	Age	ADI-social	ADI-communication	ADI-RRB	ADOS-TOTAL	ADOS-SA	ADOS-RRB
1	ASD	Adults	Nijmegen	F	71	66	70	17.49	9	14	2	7	8	1
2	ASD	IQ <70	Cambridge	M	73	66	67	24.29	10	9	5	9	8	10
3	TD	Adults	KCL	M	142	136	142	23.08	-	-	-	-	-	-
4	ASD	Adolescents	Nijmegen	M	89	64	78	12.07	17	12	1	3	5	1
5	TD	Adults	Cambridge	F	104	117	111	18.26	-	-	-	-	-	-
6	ASD	Adolescents	Cambridge	M	103	120	113	14.53	16	12	1	7	8	5
7	ASD	Adolescents	KCL	M	144	134	142	16.82	20	12	5	6	6	7
8	ASD	Children	KCL	F	98	106	102	9.31	25	20	8	8	7	9
9	ASD	Children	KCL	F	116	133	127	10.68	10	9	8	1	2	6
10	TD	Adults	Utrecht	M	105	85	96	22.06	-	-	-	-	-	-
11	TD	Children	Cambridge	M	108	129	120	8.62	-	-	-	-	-	-
12	ASD	Children	Mannheim	M	-	69	-	9.69	20	10	8	-	-	-
13	ASD	Adults	KCL	M	133	120	130	19.44	18	19	8	-	-	-
14	ASD	Adolescents	Cambridge	F	112	104	109	12.11	27	17	8	2	4	1
15	ASD	Children	Nijmegen	M	92	94	93	11.57	24	20	7	1	3	1

Finally, Supplementary Tables S6 and S7 shows the associations for the deviations from the normative model and symptom scales.

*Supplementary Table S6: Clinical relevance of the deviations; Significant correlation (Spearman) between the mean of extreme deviation in each cortical parcel and symptoms measured by ADOS and ADI scores (P_value < 0.05). * indicates the regions survived after FDR correction*

Parcel	Correlation Coefficient, r	Parcel	Correlation Coefficient, r
ADI_social, Female		***ADI_RRB, Male***	
superiortemporal	0.22(L)	superiorfrontal	0.23*(L), 0.2(R)
lateralorbitofrontal	0.22(R)	superiortemporal	0.16(L)
parsopercularis	0.22(R)	insula	0.16(L)
precuneus	0.24(R)	caudalanteriorcingulate	0.2(R)
temporalpole	0.26(R)	fusiform	0.15(R)
ADI_social, Male		paracentral	0.15(R)
superiortemporal	0.17(L)	parstriangularis	0.17(R)
lateralorbitofrontal	0.15(R)	temporalpole	0.16(R)
ADI_communication, Female		***ADOS_II_CSS, Female***	
caudalmiddlefrontal	0.21(L)	parsopercularis	0.25(L)
lateralorbitofrontal	0.25(L)	***ADOS_II_CSS, Male***	
parsopercularis	0.22(L)	entorhinal	0.18(R)
pericalcarine	0.23(L)	medialorbitofrontal	0.19(R)
postcentral	0.22(L)	temporalpole	0.16(R)
precentral	0.23(L)	***ADOS_II_Social Affect, Female***	
superiorparietal	0.23(L)	entorhinal	-0.23(L)
inferiortemporal	0.23(R)	lateraloccipital	-0.25(L)
lateraloccipital	0.30(R)	lingual	-0.22(L)
middletemporal	0.30(R)	middletemporal	-0.24(L)
precuneus	0.21(R)	parsopercularis	0.23(L)
rostralmiddlefrontal	0.21(R)	***ADOS_II_Social Affect, Male***	
ADI_communication, Male		medialorbitofrontal	0.18(R)
superiortemporal	0.17(L)	temporalpole	0.15(R)
supramarginal	0.15(L)	***ADOS_II_RRB, Female***	
insula	0.17(L)	caudalmiddlefrontal	0.33*(L), 0.31(R)
superiorparietal	0.15(R)	entorhinal	0.23(L)
paracentral	0.15(R)	fusiform	0.33*(L)
ADI_RRB, Female		inferiorparietal	0.26(L),0.25(R)
caudalmiddlefrontal	0.31*(L), 0.24(R)	inferiortemporal	0.25(L)
entorhinal	0.24(L)	lateraloccipital	0.23(L)
inferiortemporal	0.22(L), 0.28(R)	lingual	0.38*(L),0.34(R)
lateraloccipital	0.30*(L)	rostralmiddlefrontal	0.23(L),0.31(R)

Parcel	Correlation Coefficient, r	Parcel	Correlation Coefficient, r
lateralorbitofrontal	0.32*(L)	superiorfrontal	0.33*(L),0.24(R)
medialorbitofrontal	0.26(L), 0.24(R)	superiorparietal	0.25(L)
parsopercularis	0.28(L)	supramarginal	0.33(L)
parstriangularis	0.31*(L)	caudalanteriorcingulate	0.26(R)
precentral	0.25(L), 0.25(R)	middletemporal	0.24(R)
precuneus	0.22(L)	paracentral	0.25(R)
superiorparietal	0.24(L)	pericalcarine	0.26(R)
inferiorparietal	0.24(R)	precentral	0.29(R)
middletemporal	0.32*(R)	precuneus	0.29(R)
rostralmiddlefrontal	0.28(R)	frontalpole	0.32(R)
superiortemporal	0.23(R)	***ADOS_II_RRB, Male***	
supramarginal	0.36(R)	lateralorbitofrontal	0.18(L)
ADI_RRB, Male		medialorbitofrontal	0.21(L)
entorhinal	0.19(L)	rostralmiddlefrontal	0.16(L)
inferiortemporal	0.14(L), 0.17(R)	insula	0.16(L)
lingual	0.14(L)	entorhinal	0.15(R)
middletemporal	0.18(L), 0.14(R)	fusiform	0.17(R)
parahippocampal	0.14(L)	inferiortemporal	0.21(R)
rostralanteriorcingulate	0.15(L), 0.16(R)	middletemporal	0.23(R)
rostralmiddlefrontal	0.15(L)	postcentral	0.18(R)

*Supplementary Table S7: Clinical relevance of the deviations across the whole brain; Significant correlation (Spearman) between extreme value across all the regions and symptoms measured by ADOS and ADI scores (P_value < 0.05). * indicates the regions survived after FDR correction*

	ADI			ADOS		
	ADI-social	***ADI-communication***	***ADI-RRB***	***ADOS-TOTAL***	***ADOS-SA***	***ADOS-RRB***
Female	0.03	0.13	0.16	-0.02	-0.13	0.30*
Male	0.11	0.10	0.20*	0.09	0.03	0.20*

Supplementary References

1. Simonoff E, Pickles A, Charman T, Chandler S, Loucas T, Baird G (2008): Psychiatric Disorders in Children With Autism Spectrum Disorders: Prevalence, Comorbidity, and Associated Factors in a Population-Derived Sample. *J Am Acad Child Adolesc Psychiatry*. 47: 921–929.

2. Wong AYS, Hsia Y, Chan EW, Murphy DGM, Simonoff E, Buitelaar JK, Wong ICK (2014): The Variation of Psychopharmacological Prescription Rates for People With Autism Spectrum Disorder (ASD) in 30 Countries. *Autism Res*. . doi: 10.1002/aur.1391.

3. Loth E, Charman T, Mason L, Tillmann J, Jones EJH, Wooldridge C, *et al.* (2017): The EU-AIMS Longitudinal European Autism Project (LEAP): design and methodologies to identify and validate stratification biomarkers for autism spectrum disorders. *Mol Autism*. 8: 24.

4. Rasmussen CE, Williams CKI (2006): Model Selection and Adaptation of Hyperparameters. *Gaussian Process Mach Learn (Adaptive Comput Mach Learn Ser*. 105–128.

5. Rasmussen CE, Williams CKI (2006): *Gaussian Processes for Machine Learning. Adapt Comput Mach Learn*. . doi: 10.1142/S0129065704001899.

6. Marquand AF, Wolfers T, Mennes M, Buitelaar J, Beckmann CF (2016): Beyond Lumping and Splitting: A Review of Computational Approaches for Stratifying Psychiatric Disorders. *Biol Psychiatry Cogn Neurosci Neuroimaging*. 1: 433–447.

7. Dale AM, Fischl B, Sereno MI (1999): Cortical surface-based analysis: I. Segmentation and surface reconstruction. *Neuroimage*. . doi: 10.1006/nimg.1998.0395.

8. Rosen AFG, Roalf DR, Ruparel K, Blake J, Seelaus K, Villa LP, *et al.* (2018): Quantitative assessment of structural image quality. *Neuroimage*. . doi: 10.1016/j.neuroimage.2017.12.059.

9. Charman T, Loth E, Tillmann J, Crawley D, Wooldridge C, Goyard D, *et al.* (2017): The EU-AIMS Longitudinal European Autism Project (LEAP): clinical characterisation. *Mol Autism*. 8: 27.

10. Goodman R, Ford T, Richards H, Gatward R, Meltzer H (2000): The Development and Well-Being Assessment: description and initial validation of an integrated assessment of child and adolescent psychopathology. *J Child Psychol Psychiatry*. . doi: 10.1111/j.1469-7610.2000.tb02345.x.